

Adaptive-Sparse Polynomial Dimensional Decomposition Methods for High-Dimensional Stochastic Computing

Vaibhav Yadav, Sharif Rahman^{1,*}

College of Engineering, The University of Iowa, Iowa City, Iowa 52242, U.S.A.

Abstract

This article presents two novel adaptive-sparse polynomial dimensional decomposition (PDD) methods for solving high-dimensional uncertainty quantification problems in computational science and engineering. The methods entail global sensitivity analysis for retaining important PDD component functions, and a full- or sparse-grid dimension-reduction integration or quasi Monte Carlo simulation for estimating the PDD expansion coefficients. A unified algorithm, endowed with two distinct ranking schemes for grading component functions, was created for their numerical implementation. The fully adaptive-sparse PDD method is comprehensive and rigorous, leading to the second-moment statistics of a stochastic response that converges to the exact solution when the tolerances vanish. A partially adaptive-sparse PDD method, obtained through regulated adaptivity and sparsity, is economical and is, therefore, expected to solve practical problems with numerous variables. Compared with past developments, the adaptive-sparse PDD methods do not require its truncation parameter(s) to be assigned *a priori* or arbitrarily. The numerical results reveal that an adaptive-sparse PDD method achieves a desired level of accuracy with considerably fewer coefficients compared with existing PDD approximations. For a required accuracy in calculating the probabilistic response characteristics, the new bivariate adaptive-sparse PDD method is more efficient than the existing bivariate truncated PDD method by almost an order of magnitude. Finally, stochastic dynamic analysis of a disk brake system was performed, demonstrating the ability of the new methods to tackle practical engineering problems.

Keywords:

ANOVA, HDMR, PDD, stochastic dynamics, uncertainty quantification

1. Introduction

Uncertainty quantification, an emerging multidisciplinary field blending physical and mathematical sciences, characterizes the discrepancy between model-based simulations and physical reality in terms of the statistical moments, probability law, and other relevant properties of a complex system response. For practical applications, encountering a large number of input random variables is not uncommon, where an output function of interest, defined algorithmically via expensive finite-element analysis (FEA) or similar numerical calculations, is all too often expensive to evaluate. The most promising stochastic methods available today are perhaps the collocation [6, 10] and polynomial chaos expansion (PCE) [14, 39] methods, including sparse-grid techniques [18], which have found many successful applications. However, for truly high-dimensional systems, they require astronomically large numbers of terms or coefficients, succumbing to the curse of dimensionality [1]. Therefore, alternative computational methods

capable of exploiting low effective dimensions of multivariate functions, such as the polynomial dimensional decomposition (PDD) method, are most desirable. Readers, not familiar with but interested in PDD, are referred to the authors' past works [25–27, 32].

For practical applications, the PDD must be truncated with respect to S and m , where S and m define the largest degree of interactions among input variables and largest order of orthogonal polynomials, respectively, retained in a concomitant approximation. These truncation parameters depend on the dimensional structure and nonlinearity of a stochastic response. The higher the values of S and m , the higher the accuracy, but also the computational cost that is endowed with an S th- or m th-order polynomial computational complexity. However, the dimensional hierarchy or nonlinearity, in general, is not known *a priori*. Therefore, indiscriminately assigning the truncation parameters is not desirable, nor is it possible to do so when a stochastic solution is obtained via complex numerical algorithms. In which case, one must perform these truncations automatically by progressively drawing in higher-variate or higher-order contributions as appropriate. Furthermore, all S -variate component functions of PDD may not contribute equally or even appreciably to be considered in the resulting approximation. Hence, a sparse approximation, expelling component functions with negligible contributions, should be considered as

^{*}Grant sponsor: U.S. National Science Foundation; Grant Nos. CMMI-0653279, CMMI-1130147.

^{*}Corresponding author.

Email addresses: vaibhav-yadav@uiowa.edu (Vaibhav Yadav), rahman@engineering.uiowa.edu (Sharif Rahman)

¹Professor.

well.

Addressing some of the aforementioned concerns have led to adaptive versions of the cut-high-dimensional model representation (cut-HDMR) [20] and the anchored decomposition [43], employed in conjunction with the sparse-grid collocation methods, for solving stochastic problems in fluid dynamics. Several adaptive variants of the PCE [2, 19, 37] method have also appeared. It is important to clarify that the cut-HDMR and anchored decompositions are the same as the referential dimensional decomposition (RDD) [28, 30]. Therefore, both adaptive methods essentially employ RDD for multivariate function approximations, where the mean values of random input are treated as the reference or anchor point – a premise originally proposed by Xu and Rahman [41]. The developments of these adaptive methods were motivated by the fact that an RDD approximation requires only function evaluations, as opposed to high-dimensional integrals required for an ANOVA Dimensional Decomposition (ADD) approximation. However, a recent error analysis [30] reveals sub-optimality of RDD approximations, meaning that an RDD approximation, regardless of how the reference point is chosen, cannot be better than an ADD approximation for identical degrees of interaction. The analysis also finds ADD approximations to be exceedingly more precise than RDD approximations at higher-variate truncations. In addition, the criteria implemented in existing adaptive methods are predicated on retaining higher-variate component functions by examining the second-moment properties of only univariate component functions, where the largest degree of interaction and polynomial order in the approximation are still left to the user's discretion, instead of being determined automatically based on the problem being solved. Therefore, more intelligently derived adaptive-sparse approximations and decompositions rooted in ADD or PDD should be explored by developing relevant criteria and acceptable error thresholds. These enhancements, some of which are indispensable, should be pursued without sustaining significant additional cost.

This paper presents two new adaptive-sparse versions of the PDD method – the fully adaptive-sparse PDD method and a partially adaptive-sparse PDD method – for solving high-dimensional stochastic problems commonly encountered in computational science and engineering. The methods are based on (1) variance-based global sensitivity analysis for defining the pruning criteria to retain important PDD component functions; (2) a full- or sparse-grid dimension-reduction integration or quasi Monte Carlo simulation (MCS) for estimating the PDD expansion coefficients. Section 2 briefly describes existing dimensional decompositions, including PDD and its S -variate, m th-order approximation, to be contrasted with the proposed methods. Two adaptive-sparse PDD methods are formally presented in Section 3, along with a computational algorithm and a flowchart for numerical implementation of the method. Two different approaches for calculating the PDD coefficients, one emanating from dimension-reduction integration and the other employing quasi MCS, are explained in Section 4. Section 5 presents three numerical examples for probing the accuracy, efficiency, and convergence properties of the proposed methods, including a comparison with the existing PDD methods. Sec-

tion 6 reports a large-scale stochastic dynamics problem solved using a proposed adaptive-sparse method. Finally, conclusions are drawn in Section 7.

2. Dimensional Decompositions

Let \mathbb{N} , \mathbb{N}_0 , \mathbb{R} , and \mathbb{R}_0^+ represent the sets of positive integer (natural), non-negative integer, real, and non-negative real numbers, respectively. For $k \in \mathbb{N}$, denote by \mathbb{R}^k the k -dimensional Euclidean space, by \mathbb{N}_0^k the k -dimensional multi-index space, and by $\mathbb{R}^{k \times k}$ the set of $k \times k$ real-valued matrices. These standard notations will be used throughout the paper.

Let (Ω, \mathcal{F}, P) be a complete probability space, where Ω is a sample space, \mathcal{F} is a σ -field on Ω , and $P : \mathcal{F} \rightarrow [0, 1]$ is a probability measure. With \mathcal{B}^N representing the Borel σ -field on \mathbb{R}^N , $N \in \mathbb{N}$, consider an \mathbb{R}^N -valued random vector $\mathbf{X} := (X_1, \dots, X_N) : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^N, \mathcal{B}^N)$, which describes the statistical uncertainties in all system and input parameters of a high-dimensional stochastic problem. The probability law of \mathbf{X} is completely defined by its joint probability density function $f_{\mathbf{X}} : \mathbb{R}^N \rightarrow \mathbb{R}_0^+$. Assuming independent coordinates of \mathbf{X} , its joint probability density $f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^N f_i(x_i)$ is expressed by a product of marginal probability density functions f_i of X_i , $i = 1, \dots, N$, defined on the probability triple $(\Omega_i, \mathcal{F}_i, P_i)$ with a bounded or an unbounded support on \mathbb{R} . For a given $u \subseteq \{1, \dots, N\}$, $f_{\mathbf{X}_{-u}}(\mathbf{x}_{-u}) := \prod_{i=1, i \notin u}^N f_i(x_i)$ defines the marginal density function of $\mathbf{X}_{-u} := \mathbf{X}_{\{1, \dots, N\} \setminus u}$.

2.1. ANOVA Dimensional Decomposition

Let $y(\mathbf{X}) := y(X_1, \dots, X_N)$, a real-valued, measurable transformation on (Ω, \mathcal{F}) , define a stochastic response to a high-dimensional random input and $\mathcal{L}_2(\Omega, \mathcal{F}, P)$ represent a Hilbert space of square-integrable functions y with respect to the induced generic measure $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$ supported on \mathbb{R}^N . The ANOVA dimensional decomposition, expressed by the recursive form [8, 30, 35]

$$y(\mathbf{X}) = \sum_{u \subseteq \{1, \dots, N\}} y_u(\mathbf{X}_u), \quad (1)$$

$$y_0 = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad (2)$$

$$y_u(\mathbf{X}_u) = \int_{\mathbb{R}^{N-|u|}} y(\mathbf{X}_u, \mathbf{x}_{-u}) f_{\mathbf{X}_{-u}}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} - \sum_{v \subset u} y_v(\mathbf{X}_v), \quad (3)$$

is a finite, hierarchical expansion in terms of its input variables with increasing dimensions, where $u \subseteq \{1, \dots, N\}$ is a subset with the complementary set $-u = \{1, \dots, N\} \setminus u$ and cardinality $0 \leq |u| \leq N$, and y_u is a $|u|$ -variate component function describing a constant or the interactive effect of $\mathbf{X}_u = (X_{i_1}, \dots, X_{i_{|u|}})$, $1 \leq i_1 < \dots < i_{|u|} \leq N$, a subvector of \mathbf{X} , on y when $|u| = 0$ or $|u| > 0$. The summation in Equation (1) comprises 2^N terms, with each term depending on a group of variables indexed by a particular subset of $\{1, \dots, N\}$, including the empty set \emptyset .

The ADD component functions y_u , $u \subseteq \{1, \dots, N\}$, have two remarkable properties: (1) the component functions, y_u , $\emptyset \neq u \subseteq \{1, \dots, N\}$, have zero means; and (2) two distinct component functions y_u and y_v , where $u \subseteq \{1, \dots, N\}$,

$v \subseteq \{1, \dots, N\}$, and $u \neq v$, are orthogonal [30]. However, the ADD component functions are difficult to obtain, because they require calculation of high-dimensional integrals.

2.2. Referential Dimensional Decomposition

Consider a reference point $\mathbf{c} = (c_1, \dots, c_N) \in \mathbb{R}^N$ and the associated Dirac measure $\prod_{i=1}^N \delta(x_i - c_i)dx_i$. The referential dimensional decomposition is created when $\prod_{i=1}^N \delta(x_i - c_i)dx_i$ replaces the probability measure in Equations (1)-(3), leading to the recursive form

$$y(\mathbf{X}) = \sum_{u \subseteq \{1, \dots, N\}} w_u(\mathbf{X}_u; \mathbf{c}), \quad (4)$$

$$w_\emptyset = y(\mathbf{c}), \quad (5)$$

$$w_u(\mathbf{X}_u; \mathbf{c}) = y(\mathbf{X}_u, \mathbf{c}_{-u}) - \sum_{v \subset u} w_v(\mathbf{X}_v; \mathbf{c}), \quad (6)$$

also known as cut-HDMR [23], anchored decomposition [17], and anchored-ANOVA decomposition [15], with the latter two referring to the reference point as the anchor. Xu and Rahman introduced Equations (4)-(6) with the aid of Taylor series expansion, calling them dimension-reduction [40] and decomposition [41] methods for statistical moment and reliability analyses, respectively, of mechanical systems. Compared with ADD, RDD lacks orthogonal features, but its component functions are easier to obtain as they only involve function evaluations at a chosen reference point.

2.3. Polynomial Dimensional Decomposition

Let $\{\psi_{ij}(X_i); j = 0, 1, \dots\}$ be a set of orthonormal polynomial basis functions in the Hilbert space $\mathcal{L}_2(\Omega_i, \mathcal{F}_i, P_i)$ that is consistent with the probability measure P_i of X_i , where $i = 1, \dots, N$. For a given $\emptyset \neq u = \{i_1, \dots, i_{|u|}\} \subseteq \{1, \dots, N\}$, $1 \leq |u| \leq N$, $1 \leq i_1 < \dots < i_{|u|} \leq N$, denote a product probability triple by $(\times_{p=1}^{|u|} \Omega_{i_p}, \times_{p=1}^{|u|} \mathcal{F}_{i_p}, \times_{p=1}^{|u|} P_{i_p})$, and the associated space of square integrable $|u|$ -variate component functions of y by $\mathcal{L}_2(\times_{p=1}^{|u|} \Omega_{i_p}, \times_{p=1}^{|u|} \mathcal{F}_{i_p}, \times_{p=1}^{|u|} P_{i_p}) := \{y_u : \int_{\mathbb{R}^{|u|}} y_u^2(\mathbf{x}_u) f_{\mathbf{x}_u}(\mathbf{x}_u) d\mathbf{x}_u < \infty\}$, which is a Hilbert space. Since the joint density of $(X_{i_1}, \dots, X_{i_{|u|}})$ is separable (independence), i.e., $f_{\mathbf{x}_u}(\mathbf{x}_u) = \prod_{p=1}^{|u|} f_{i_p}(x_{i_p})$, the product polynomial $\psi_{\mathbf{j}_{|u|}}(\mathbf{X}_u) := \prod_{p=1}^{|u|} \psi_{i_p j_p}(X_{i_p})$, where $\mathbf{j}_{|u|} = (j_1, \dots, j_{|u|}) \in \mathbb{N}_0^{|u|}$, a $|u|$ -dimensional multi-index with ∞ -norm $\|\mathbf{j}_{|u|}\|_\infty := \max(j_1, \dots, j_{|u|})$, constitutes an orthonormal basis in $\mathcal{L}_2(\times_{p=1}^{|u|} \Omega_{i_p}, \times_{p=1}^{|u|} \mathcal{F}_{i_p}, \times_{p=1}^{|u|} P_{i_p})$.

The orthogonal polynomial expansion of a non-constant $|u|$ -variate component function becomes [25, 26]

$$y_u(\mathbf{X}_u) = \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|} \\ j_1, \dots, j_{|u|} \neq 0}} C_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \quad \emptyset \neq u \subseteq \{1, \dots, N\}, \quad (7)$$

with

$$C_{u\mathbf{j}_{|u|}} := \int_{\mathbb{R}^N} y(\mathbf{x}) \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}, \quad \emptyset \neq u \subseteq \{1, \dots, N\}, \quad \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, \quad (8)$$

representing the corresponding expansion coefficient. The end result of combining Equations (1) and (7) is the PDD [25, 26],

$$y(\mathbf{X}) = y_\emptyset + \sum_{\emptyset \neq u \subseteq \{1, \dots, N\}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|} \\ j_1, \dots, j_{|u|} \neq 0}} C_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \quad (9)$$

providing an exact, hierarchical expansion of y in terms of an infinite number of coefficients or orthonormal polynomials. All component functions y_u , $\emptyset \neq u \subseteq \{1, \dots, N\}$, in Equation (7) have zero means and satisfy the orthogonal properties of the ADD. Therefore, PDD can be viewed as the polynomial version of ADD, inheriting all desirable properties of ADD.

2.4. Truncated Dimensional Decompositions

The three dimensional decompositions – ADD, RDD and PDD – are grounded on a fundamental conjecture known to be true in many real-world applications: given a high-dimensional function y , its $|u|$ -variate component functions decay rapidly with respect to $|u|$, leading to accurate lower-variate approximations of y . Indeed, given the integers $0 \leq S < N$ and $1 \leq m < \infty$ for all $1 \leq |u| \leq S$, the truncated dimensional decompositions

$$\tilde{y}_S(\mathbf{X}) = \sum_{\substack{u \subseteq \{1, \dots, N\} \\ 0 \leq |u| \leq S}} y_u(\mathbf{X}_u), \quad (10)$$

$$\hat{y}_S(\mathbf{X}; \mathbf{c}) = \sum_{\substack{u \subseteq \{1, \dots, N\} \\ 0 \leq |u| \leq S}} w_u(\mathbf{X}_u; \mathbf{c}), \quad (11)$$

$$\tilde{y}_{S,m}(\mathbf{X}) = y_\emptyset + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|} \\ \|\mathbf{j}_{|u|}\|_\infty \leq m \\ j_1, \dots, j_{|u|} \neq 0}} C_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \quad (12)$$

respectively, describe S -variate ADD, RDD, and PDD approximations, which for $S > 0$ include interactive effects of at most S input variables X_{i_1}, \dots, X_{i_S} , $1 \leq i_1 < \dots < i_S \leq N$, on y . It is elementary to show that when $S \rightarrow N$ and/or $m \rightarrow \infty$, \tilde{y}_S , \hat{y}_S , and $\tilde{y}_{S,m}$ converge to y in the mean-square sense, generating a hierarchical and convergent sequence of approximation of y from each decomposition.

2.4.1. ADD and RDD Errors

For ADD or RDD to be useful, what are the approximation errors committed by $\tilde{y}_S(\mathbf{X})$ and $\hat{y}_S(\mathbf{X}; \mathbf{c})$ in Equations (10) and (11)? More importantly, for a given $0 \leq S < N$, which approximation between ADD and RDD is better? Since the RDD approximation depends on the reference point \mathbf{c} , no analytical error analysis is possible if \mathbf{c} is deterministic or arbitrarily chosen. However, if \mathbf{c} follows the same probability measure of \mathbf{X} , then the error committed by an RDD approximation on average can be compared with the error from an ADD approximation, as follows.

Theorem 1. Let $\mathbf{c} = (c_1, \dots, c_N) \in \mathbb{R}^N$ be a random vector with the joint probability density function of the form $f_{\mathbf{x}}(\mathbf{c}) =$

$\prod_{j=1}^N f_j(c_j)$, where f_j is the marginal probability density function of the j th coordinate of $\mathbf{X} = (X_1, \dots, X_N)$. Define two second-moment errors

$$e_{S,A} := \mathbb{E}[(y(\mathbf{X}) - \tilde{y}_S(\mathbf{X}))^2] := \int_{\mathbb{R}^N} [y(\mathbf{x}) - \tilde{y}_S(\mathbf{x})]^2 f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad (13)$$

and

$$\begin{aligned} e_{S,R}(\mathbf{c}) &:= \mathbb{E}[(y(\mathbf{X}) - \hat{y}_S(\mathbf{X}; \mathbf{c}))^2] \\ &:= \int_{\mathbb{R}^N} [y(\mathbf{x}) - \hat{y}_S(\mathbf{x}; \mathbf{c})]^2 f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \end{aligned} \quad (14)$$

committed by the S -variate ADD and RDD approximations, respectively, of $y(\mathbf{X})$. Then the lower and upper bounds of the expected error $\mathbb{E}[e_{S,R}] := \int_{\mathbb{R}^N} e_{S,R}(\mathbf{c}) f_{\mathbf{X}}(\mathbf{c}) d\mathbf{c}$ from the S -variate RDD approximation, expressed in terms of the error $e_{S,A}$ from the S -variate ADD approximation, are

$$2^{S+1} e_{S,A} \leq \mathbb{E}[e_{S,R}] \leq \left[1 + \sum_{k=0}^S \binom{N-S+k-1}{k} \binom{N}{S-k} \right] e_{S,A}, \quad (15)$$

where $0 \leq S < N$, $S+1 \leq N < \infty$.

Proof. See Theorem 4.12 and Corollary 4.13 of Rahman [30]. \square

Remark 1. Theorem 1 reveals that the expected error from the univariate ($S = 1$) RDD approximation is at least four times larger than the error from the univariate ADD approximation. In contrast, the expected error from the bivariate ($S = 2$) RDD approximation can be eight or more times larger than the error from the bivariate ADD approximation. Given an arbitrary truncation, an ADD approximation is superior to an RDD approximation. In addition, the RDD approximations may perpetrate very large errors at upper bounds when there exist a large number of variables and appropriate conditions. Therefore, existing adaptive methods [20, 43] anchored in RDD approximations should be used with caveat. Furthermore, the authors advocate using PDD for adaptivity, but doing so engenders its own computational challenges, to be explained in the forthcoming sections.

2.4.2. Statistical Moments of PDD

Applying the expectation operator on $\tilde{y}_{S,m}(\mathbf{X})$ and $(\tilde{y}_{S,m}(\mathbf{X}) - y_0)^2$ and noting the zero-mean and orthogonal properties of PDD component functions, the mean [27]

$$\mathbb{E}[\tilde{y}_{S,m}(\mathbf{X})] = y_0 \quad (16)$$

of the S -variate, m th-order PDD approximation matches the exact mean $\mathbb{E}[y(\mathbf{X})]$, regardless of S or m , and the approximate variance [27]

$$\begin{aligned} \sigma_{S,m}^2 &:= \mathbb{E}[(\tilde{y}_{S,m}(\mathbf{X}) - \mathbb{E}[\tilde{y}_{S,m}(\mathbf{X})])^2] \\ &= \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, \|\mathbf{j}_{|u|}\|_{\infty} \leq m \\ j_1, \dots, j_{|u|} \neq 0}} C_{u\mathbf{j}_{|u|}}^2 \end{aligned} \quad (17)$$

is calculated as the sum of squares of the expansion coefficients from the S -variate, m th-order PDD approximation of $y(\mathbf{X})$. Clearly, the approximate variance approaches the exact variance [27]

$$\sigma^2 := \mathbb{E}[(y(\mathbf{X}) - \mathbb{E}[y(\mathbf{X})])^2] = \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|} \\ j_1, \dots, j_{|u|} \neq 0}} C_{u\mathbf{j}_{|u|}}^2 \quad (18)$$

of y when $S \rightarrow N$ and $m \rightarrow \infty$. The mean-square convergence of $\tilde{y}_{S,m}$ is guaranteed as y , and its component functions are all members of the associated Hilbert spaces.

The S -variate, m th-order PDD approximation $\tilde{y}_{S,m}(\mathbf{X})$ in Equation (12) contains

$$\tilde{K}_{S,m} = 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, \|\mathbf{j}_{|u|}\|_{\infty} \leq m \\ j_1, \dots, j_{|u|} \neq 0}} 1 = \sum_{k=0}^S \binom{N}{k} m^k \quad (19)$$

number of PDD coefficients and corresponding orthonormal polynomials. Therefore, the computational complexity of a truncated PDD is polynomial, as opposed to exponential, thereby alleviating the curse of dimensionality to some extent.

Remark 2. Constructing a PDD approximation by pre-selecting S and/or m , unless they are quite small, is computationally intensive, if not impossible, for high-dimensional uncertainty quantification. In other words, the existing PDD is neither scalable nor adaptable, which is crucial for solving industrial-scale stochastic problems. A requisite theoretical basis and innovative numerical algorithms for overcoming these limitations are presented in Section 3.

Remark 3. The PDD approximation and its second-moment analysis require the expansion coefficients $C_{u\mathbf{j}_{|u|}}$, which, according to their definition in Equation (8), involve various N -dimensional integrals over \mathbb{R}^N . For large N , a full numerical integration employing an N -dimensional tensor product of a univariate quadrature formula is computationally prohibitive. This is one drawback of ADD and PDD, since their component functions entail calculating high-dimensional integrals. Therefore, novel dimension-reduction integration schemes or sampling techniques, to be described in Section 4, are needed to estimate the coefficients efficiently.

2.4.3. PDD versus PCE Approximations

The long form of an S -variate, m th-order PDD approximation of $y(\mathbf{X})$ is the expansion [25, 26]

$$\begin{aligned} \tilde{y}_{S,m}(\mathbf{X}) &:= y_0 + \sum_{i=1}^N \sum_{j=1}^m C_{ij} \psi_{ij}(X_i) + \\ &\quad \sum_{i_1=1}^{N-1} \sum_{i_2=i_1+1}^N \sum_{j_2=1}^m \sum_{j_1=1}^m C_{i_1 i_2 j_1 j_2} \psi_{i_1 j_1}(X_{i_1}) \psi_{i_2 j_2}(X_{i_2}) + \dots + \\ &\quad \underbrace{\sum_{i_1=1}^{N-s+1} \dots \sum_{i_s=i_{s-1}+1}^N}_{S \text{ sums}} \underbrace{\sum_{j_1=1}^m \dots \sum_{j_s=1}^m}_{S \text{ sums}} C_{i_1 \dots i_s j_1 \dots j_s} \prod_{q=1}^S \psi_{i_q j_q}(X_{i_q}) \end{aligned}$$

(20)

in terms of random orthonormal polynomials $\psi_{ij}(X_i)$, $i = 1, \dots, N$, $j = 1, \dots, m$, of input variables X_1, \dots, X_N with increasing dimensions, where y_0 and $C_{i_1 \dots i_S j_1 \dots j_S}$, $1 \leq i_1 < \dots < i_S \leq N$, $j_1, \dots, j_S = 1, \dots, m$, are the PDD expansion coefficients. In contrast, a p th-order PCE approximation of $y(\mathbf{X})$, where $0 \leq p < \infty$, has the representation [14]

$$\begin{aligned} \tilde{y}_p(\mathbf{X}) := & a_0 \Gamma_0 + \sum_{i=1}^N a_i \Gamma_1(X_i) + \sum_{i_1=1}^N \sum_{i_2=1}^N a_{i_1 i_2} \Gamma_2(X_{i_1}, X_{i_2}) \\ & + \dots + \sum_{i_1=1}^N \dots \sum_{i_p=1}^N a_{i_1 \dots i_p} \Gamma_p(X_{i_1}, \dots, X_{i_p}) \end{aligned} \quad (21)$$

in terms of random polynomial chaoses $\Gamma_p(X_{i_1}, \dots, X_{i_p})$, $1 \leq i_1 \leq \dots \leq i_p \leq N$, of input variables X_{i_1}, \dots, X_{i_p} with increasing orders, where a_0 and $a_{i_1 \dots i_p}$ are the PCE expansion coefficients. The polynomial chaoses are various combinations of tensor products of sets of univariate orthonormal polynomials. Therefore, both expansions share the same orthonormal polynomials, and their coefficients require evaluating similar high-dimensional integrals.

Remark 4. The PDD and PCE when truncated are not the same. In fact, two important observations jump out readily. First, the terms in the PCE approximation are organized with respect to the order of polynomials. In contrast, the PDD approximation is structured with respect to the degree of interaction between a finite number of random variables. Therefore, significant differences may exist regarding the accuracy, efficiency, and convergence properties of their truncated sum or series. Second, if a stochastic response is highly nonlinear, but contains rapidly diminishing interactive effects of multiple random variables, the PDD approximation is expected to be more effective than the PCE approximation. This is because the lower-variate (univariate, bivariate, *etc.*) terms of the PDD approximation can be just as nonlinear by selecting appropriate values of m in Equation (20). In contrast, many more terms and expansion coefficients are required to be included in the PCE approximation to capture such high nonlinearity.

In reference to a past study [32], consider two mean-squared errors, $e_{S,m} := \mathbb{E}[y(\mathbf{X}) - \tilde{y}_{S,m}(\mathbf{X})]^2$ and $e_p := \mathbb{E}[y(\mathbf{X}) - \tilde{y}_p(\mathbf{X})]^2$, owing to the S -variate, m th-order PDD approximation $\tilde{y}_{S,m}(\mathbf{X})$ and p th-order PCE approximation $\tilde{y}_p(\mathbf{X})$, respectively, of $y(\mathbf{X})$. For a class of problems where the interactive effects of S input variables on a stochastic response get progressively weaker as $S \rightarrow N$, then the PDD and PCE errors for identical expansion orders can be weighed against each other. For this special case, set $m = p$ and assume that $C_{i_1 \dots i_S j_1 \dots j_S} = 0$, where $s = S + 1, \dots, N$, $1 \leq i_1 < \dots < i_s \leq N$, $j_1, \dots, j_s = 1, \dots, \infty$. Then it can be shown that $e_m \geq e_{S,m}$, demonstrating larger error from the PCE approximation than from the PDD approximation [32]. In the limit, when $S = N$, $e_m \geq e_{N,m}$, regardless of the values of the expansions coefficients. In other words, the

N -variate, m th-order PDD approximation cannot be worse than the m th-order PCE approximation. When $S < N$ and $C_{i_1 \dots i_S j_1 \dots j_S}$, $s = S + 1, \dots, N$, $1 \leq i_1 < \dots < i_s \leq N$, $j_1, \dots, j_s = 1, \dots, \infty$, are not negligible and arbitrary, numerical convergence analysis is required for comparing these two errors. Indeed, numerical analyses of mathematical functions or simple dynamic systems reveal markedly higher convergence rates of the PDD approximation than the PCE approximation [32]. From the comparison of computational efforts, required to estimate with the same precision the frequency distributions of complex dynamic systems, the PDD approximation can be significantly more efficient than the PCE approximation [32].

3. Proposed Adaptive-Sparse PDD Methods

3.1. Global Sensitivity Indices

The global sensitivity analysis quantifies how an output function of interest is influenced by individual or subsets of input variables, illuminating the dimensional structure lurking behind a complex response. Indeed, these sensitivity indices have been used to rank variables, fix unessential variables, and reduce dimensions of large-scale problems [29, 36]. The authors propose to exploit these indices, developed in conjunction with PDD, for adaptive-sparse PDD approximations as follows.

The global sensitivity index of $y(\mathbf{X})$ for a subset \mathbf{X}_u , $\emptyset \neq u \subseteq \{1, \dots, N\}$, of input variables \mathbf{X} , denoted by G_u , is defined as the non-negative ratio [29, 36]

$$G_u := \frac{\mathbb{E}[y_u^2(\mathbf{X})]}{\sigma^2}, 0 < \sigma^2 < \infty, \quad (22)$$

representing the fraction of the variance of $y(\mathbf{X})$ contributed by the ADD component function y_u . Since $\emptyset \neq u \subseteq \{1, \dots, N\}$, there exist $2^N - 1$ such indices, adding up to $\sum_{u \subseteq \{1, \dots, N\}} G_u = 1$. Applying the Fourier-polynomial approximation of $y_u(\mathbf{X})$, that is, Equation (7), and noting the properties of orthonormal polynomials, the component variance

$$\mathbb{E}[y_u^2(\mathbf{X}_u)] = \sum_{\substack{\mathbf{j}_u \in \mathbb{N}_0^{|\mathbf{j}_u|} \\ j_1, \dots, j_{|\mathbf{j}_u|} \neq 0}} C_{u, \mathbf{j}_u}^2 \quad (23)$$

of y_u is the sum of squares of its PDD expansion coefficients. When the right side of Equation (23) is truncated at $\|\mathbf{j}_u\|_\infty = m_u$, where $m_u \in \mathbb{N}$, and then used to replace the numerator of Equation (22), the result is an m_u th-order approximation

$$\tilde{G}_{u, m_u} := \frac{1}{\sigma^2} \sum_{\substack{\mathbf{j}_u \in \mathbb{N}_0^{|\mathbf{j}_u|} \\ \|\mathbf{j}_u\|_\infty \leq m_u \\ j_1, \dots, j_{|\mathbf{j}_u|} \neq 0}} C_{u, \mathbf{j}_u}^2, \quad (24)$$

which approaches G_u as $m_u \rightarrow \infty$. Given $2 \leq m_u < \infty$, consider two approximate global sensitivity indices \tilde{G}_{u, m_u-1} and \tilde{G}_{u, m_u} for \mathbf{X}_u such that $\tilde{G}_{u, m_u-1} \neq 0$. Then the normalized index, defined by

$$\Delta \tilde{G}_{u, m_u} := \frac{\tilde{G}_{u, m_u} - \tilde{G}_{u, m_u-1}}{\tilde{G}_{u, m_u-1}}, \tilde{G}_{u, m_u-1} \neq 0, \quad (25)$$

represents the relative change in the approximate global sensitivity index when the largest polynomial order increases from $m_u - 1$ to m_u . The sensitivity indices \tilde{G}_{u,m_u} and $\Delta\tilde{G}_{u,m_u}$ provide an effective means to truncate the PDD in Equation (9) both adaptively and sparsely.

3.2. The Fully Adaptive-Sparse PDD Method

Let $\epsilon_1 \geq 0$ and $\epsilon_2 \geq 0$ denote two non-negative error tolerances that specify the minimum values of \tilde{G}_{u,m_u} and $\Delta\tilde{G}_{u,m_u}$, respectively. Then a fully adaptive-sparse PDD approximation

$$\bar{y}(\mathbf{X}) := y_0 + \sum_{\emptyset \neq u \subseteq \{1, \dots, N\}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_u\|_{\infty} = m_u, j_1, \dots, j_{|u|} \neq 0 \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta\tilde{G}_{u,m_u} > \epsilon_2}} C_{u\mathbf{j}_u} \psi_{u\mathbf{j}_u}(\mathbf{X}_u) \quad (26)$$

of $y(\mathbf{X})$ is formed by the subset of PDD component functions, satisfying two inclusion criteria: (1) $\tilde{G}_{u,m_u} > \epsilon_1$, and (2) $\Delta\tilde{G}_{u,m_u} > \epsilon_2$ for all $1 \leq |u| \leq N$ and $1 \leq m_u < \infty$. The first criterion requires the contribution of an m_u -th order polynomial approximation of $y_u(\mathbf{X})$ towards the variance of $y(\mathbf{X})$ to exceed ϵ_1 in order to be accommodated within the resultant truncation. The second criterion identifies the augmentation in the variance contribution from $y_u(\mathbf{X}_u)$ evoked by a single increment in the polynomial order m_u and determines if it surpasses ϵ_2 . In other words, these two criteria ascertain which interactive effects between two or more input random variables are retained and dictate the largest order of polynomials in a component function, formulating a fully adaptive-sparse PDD approximation.

When compared with the PDD in Equation (9), the adaptive-sparse PDD approximation in Equation (26) filters out the relatively insignificant component functions with a scant compromise on the accuracy of the resulting approximation. Furthermore, there is no need to pre-select the truncation parameters of the existing PDD approximation. The level of accuracy achieved by the fully adaptive-sparse PDD is meticulously controlled through the tolerances ϵ_1 and ϵ_2 . The lower the tolerance values, the higher the accuracy of the approximation. It is elementary to show that the mean-squared error in the fully adaptive-sparse PDD approximation disappears when the tolerances vanish, that is, $\bar{y}(\mathbf{X})$ approaches $y(\mathbf{X})$ as $\epsilon_1 \rightarrow 0$, $\epsilon_2 \rightarrow 0$.

3.3. A Partially Adaptive-Sparse PDD Method

Based on the authors' past experience, an S -variate PDD approximation, where $S \ll N$, is adequate, when solving real-world engineering problems, with the computational cost varying polynomially (S -order) with respect to the number of variables [25, 26]. As an example, consider the selection of $S = 2$ for solving a stochastic problem in 100 dimensions by a bivariate PDD approximation, comprising $100 \times 99/2 = 4950$ bivariate component functions. If all such component functions are included, then the computational effort for even a full bivariate PDD approximation may exceed the computational budget allocated to solving this problem. But many of these component functions contribute little to the probabilistic characteristics sought and can be safely ignored. Similar conditions may

prevail for higher-variate component functions. Henceforth, define an S -variate, partially adaptive-sparse PDD approximation

$$\bar{y}_S(\mathbf{X}) := y_0 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_u\|_{\infty} = m_u, j_1, \dots, j_{|u|} \neq 0 \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta\tilde{G}_{u,m_u} > \epsilon_2}} C_{u\mathbf{j}_u} \psi_{u\mathbf{j}_u}(\mathbf{X}_u) \quad (27)$$

of $y(\mathbf{X})$, which is attained by subsuming at most S -variate component functions, but fulfilling two relaxed inclusion criteria: (1) $\tilde{G}_{u,m_u} > \epsilon_1$ for $1 \leq |u| \leq S \leq N$, and (2) $\Delta\tilde{G}_{u,m_u} > \epsilon_2$ for $1 \leq |u| \leq S \leq N$. Again, the same two criteria are used for the degree of interaction and the order of orthogonal polynomial, but the truncations are restricted to at most S -variate component functions of y .

An S -variate, partially adaptive-sparse PDD approximation behaves differently from the S -variate, m th-order PDD approximation. While the latter approximation includes a sum containing at most S -variate component functions, the former approximation may or may not include all such component functions, depending on the tolerance ϵ_1 . For $\epsilon_1 > 0$, an S -variate, partially adaptive-sparse PDD will again trim the component functions with meager contributions. However, unlike \bar{y} converging to y , \bar{y}_S converges to the S -variate ADD approximation \tilde{y}_S , when $\epsilon_1 \rightarrow 0$, $\epsilon_2 \rightarrow 0$. If $S = N$, then both partially and fully adaptive-sparse PDD approximations coincide for identical tolerances.

As $S \rightarrow N$, $\bar{y}_S(\mathbf{X}) \rightarrow y(\mathbf{X})$ in the mean square sense. Given a rate at which $\sigma_u^2 := \mathbb{E}[y_u^2(\mathbf{X}_u)]$, the variance of an $|u|$ -variate ADD component function, decreases with $|u|$, what can be inferred on how fast $\bar{y}_S(\mathbf{X})$ converges to $y(\mathbf{X})$? Proposition 1 and subsequent discussions provide some insights.

Proposition 1. *If the variance of a zero-mean ADD component function y_u diminishes according to $\sigma_u^2 \leq cq^{-|u|}$, where $\emptyset \neq u \subseteq \{1, \dots, N\}$, and $c > 0$ and $q > 1$ are two real-valued constants, then the mean-squared error committed by $\bar{y}_S(\mathbf{X})$, $0 \leq S \leq N$, is*

$$\tilde{\epsilon}_S := \mathbb{E}[y(\mathbf{X}) - \bar{y}_S(\mathbf{X})]^2 \leq c \sum_{s=S+1}^N \binom{N}{s} q^{-s}. \quad (28)$$

Proof. The result of Proposition 1 follows by substituting the expressions of $y(\mathbf{X})$ and $\bar{y}_S(\mathbf{X})$ from Equations (1) and (10), and then using $\sigma_u^2 := \mathbb{E}[y_u^2(\mathbf{X}_u)] \leq cq^{-|u|}$. \square

When the equality holds, $\tilde{\epsilon}_S$ decays strictly monotonically with respect to S for any rate parameter q . The higher the value of S , the faster $\bar{y}_S(\mathbf{X})$ converges to $y(\mathbf{X})$ in the mean-square sense.

3.4. Stochastic Solutions

3.4.1. Second-Moment Properties

Applying the expectation operator on $\bar{y}(\mathbf{X})$ and $\bar{y}_S(\mathbf{X})$ and recognizing the zero-mean and orthogonal properties of PDD component functions, the means

$$\mathbb{E}[\bar{y}(\mathbf{X})] = \mathbb{E}[\bar{y}_S(\mathbf{X})] = y_0 \quad (29)$$

of fully and partially adaptive-sparse PDD approximations both also agree with the exact mean $\mathbb{E}[y(\mathbf{X})] = y_0$ for any ϵ_1, ϵ_2 , and S . However, the respective variances, obtained by applying the expectation operator on $(\bar{y}(\mathbf{X}) - y_0)^2$ and $(\bar{y}_S(\mathbf{X}) - y_0)^2$, vary according to

$$\begin{aligned}\bar{\sigma}^2 &:= \mathbb{E}[(\bar{y}(\mathbf{X}) - \mathbb{E}[\bar{y}(\mathbf{X})])^2] \\ &= \sum_{\emptyset \neq u \subseteq \{1, \dots, N\}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_u\|_{\infty}=m_u, j_1, \dots, j_{|u|} \neq 0 \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2}} C_{u, \mathbf{j}_u}^2\end{aligned}\quad (30)$$

and

$$\begin{aligned}\bar{\sigma}_S^2 &:= \mathbb{E}[(\bar{y}_S(\mathbf{X}) - \mathbb{E}[\bar{y}_S(\mathbf{X})])^2] \\ &= \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_u\|_{\infty}=m_u, j_1, \dots, j_{|u|} \neq 0 \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2}} C_{u, \mathbf{j}_u}^2,\end{aligned}\quad (31)$$

where the squares of the expansion coefficients are summed following the same two pruning criteria discussed in the preceding subsections. Equations (29)-(31) provide closed-form expressions of the approximate second-moment properties of any square-integrable function y in terms of the PDD expansion coefficients.

When $\epsilon_1 = \epsilon_2 = 0$, the right sides of Equations (30) and (18) coincide, whereas the right side of Equation (31) approaches that of Equation (17) for $m \rightarrow \infty$. As a consequence, the variance from the fully adaptive-sparse PDD approximation $\bar{y}(\mathbf{X})$ converges to the exact variance of $y(\mathbf{X})$ as $\epsilon_1 \rightarrow 0$ and $\epsilon_2 \rightarrow 0$. In contrast, the variance from the S -variate, partially adaptive-sparse PDD approximation $\bar{y}_S(\mathbf{X})$ does not follow suit, as it converges to the variance of the S -variate, m th-order PDD approximation $\tilde{y}_{S,m}(\mathbf{X})$ as $\epsilon_1 \rightarrow 0$ and $\epsilon_2 \rightarrow 0$, provided that $m \rightarrow \infty$. Therefore, the fully adaptive-sparse PDD approximation is more rigorous than a partially adaptive-sparse PDD approximation, but the latter can be more useful than the former when solving practical engineering problems and will be demonstrated in the Numerical Examples and Application sections.

3.4.2. Probability Distribution

Although the PDD approximations are mean-square convergent, Equations (26) and (27) can also be used to estimate higher-order moments and probability distributions, including rare-event probabilities, of sufficiently smooth stochastic responses. In this paper, the probability distribution of $y(\mathbf{X})$ was approximated by performing Monte Carlo simulation of $\bar{y}(\mathbf{X})$ and/or $\bar{y}_S(\mathbf{X})$. This simulation of the PDD approximation should not be confused with *crude* Monte Carlo simulation. The crude Monte Carlo method, which commonly requires numerical calculations of y for input samples can be expensive or even prohibitive, particularly when the sample size needs to be very large for estimating small failure probabilities. In contrast, the Monte Carlo simulation embedded in a PDD approximation requires evaluations of simple analytical functions. Therefore,

an arbitrarily large sample size can be accommodated in the PDD approximation.

It is also possible to estimate the probability distribution of $y(\mathbf{X})$ from the knowledge of the cumulant generating function of a PDD approximation, provided that it exists, and then exploit the saddle point approximation for obtaining an exponential family of approximate distributions. Readers interested in this alternative approach are referred to the authors' ongoing work on stochastic sensitivity analysis [31].

It is important to emphasize that the two truncation criteria proposed are strictly based on variance as a measure of output uncertainty. They are highly relevant when the second-moment properties of complex response is desired. For higher-order moments or rare-event probabilities, it is possible to develop alternative sensitivity indices and related pruning criteria. They are not considered here.

3.5. Numerical Implementation

The application of fully and partially adaptive-sparse PDD approximations described by Equations (26) and (27) requires selecting PDD component functions $y_u(\mathbf{X}_u)$, $\emptyset \neq u \subseteq \{1, \dots, N\}$ and assigning largest orders of their orthogonal polynomial expansions $1 \leq m_u < \infty$ efficiently such that $\tilde{G}_{u,m_u} > \epsilon_1$ and $\Delta \tilde{G}_{u,m_u} > \epsilon_2$. This section presents a unified computational algorithm and an associated flowchart developed to accomplish numerical implementation of the two proposed methods.

3.5.1. A Unified Algorithm

The iterative process for constructing an adaptive-sparse PDD approximation, whether full or partial, comprises two main stages: (1) continue incrementing the polynomial order m_u for a chosen component function $y_u(\mathbf{X}_u)$ unless the criterion $\Delta \tilde{G}_{u,m_u} > \epsilon_2$ fails; and (2) continue selecting the component functions $y_u(\mathbf{X}_u)$, $\emptyset \neq u \subseteq \{1, \dots, N\}$, unless the criterion $\tilde{G}_{u,m_u} > \epsilon_1$ fails. These two stages are first executed over all univariate PDD component functions $y_u(\mathbf{X}_u)$, $|u| = 1$, before progressing to all bivariate component functions $y_u(\mathbf{X}_u)$, $|u| = 2$, and so on, until $|u| = N$ for the fully adaptive-sparse PDD approximation or until $|u| = S$ for a partially adaptive-sparse PDD approximation, where S is specified by the user. The implementation details of the iterative process is described in Algorithm 1 and through the flowchart in Figure 1.

The first stage of the algorithm presented is predicated on accurate calculations of the sensitivity indices \tilde{G}_{u,m_u} and $\Delta \tilde{G}_{u,m_u}$, which require the variance σ^2 of $y(\mathbf{X})$ as noted by Equations (24) and (25). Since there exist an infinite number of expansion coefficients emanating from all PDD component functions, calculating the variance exactly from Equation (18) is impossible. To overcome this quandary, the authors propose to estimate the variance by utilizing all PDD expansion coefficients available at a juncture of the iterative process. For instance, let $v \in V$ be an element of the index set $V \subseteq \{1, \dots, N\}$, which comprises the subsets of $\{1, \dots, N\}$ selected so far at a given step of the iterative process. Then the approximate variance

$$\tilde{\sigma}_V^2 = \sum_{\emptyset \neq v \in V \subseteq \{1, \dots, N\}} \sum_{\substack{\mathbf{j}_v \in \mathbb{N}_0^{|v|}, \|\mathbf{j}_v\|_{\infty} \leq m_v \\ j_1, \dots, j_{|v|} \neq 0}} C_{v, \mathbf{j}_v}^2 \quad (32)$$

Algorithm 1 Adaptive-sparse polynomial dimensional decomposition.

Define S ▷ [$S \leftarrow N$ for Fully adaptive]
Define $\epsilon_1, \epsilon_2, \epsilon_3$
for $|u| \leftarrow 1$ to S **do**
 $|v| \leftarrow |u|, v \subseteq \{1, \dots, N\}$
 $m_v \leftarrow 0$
 repeat ▷ [continue incrementing the polynomial order m_v unless the ranking of component functions $y_v(\mathbf{x}_v)$ converges]
 $m_v \leftarrow m_v + 1$ ▷ [start with the polynomial order $m_v = 1$]
 calculate $C_{v,j_v}, \mathbf{j}_{|v|} \in \mathbb{N}_0^{|v|}, \|\mathbf{j}_{|v|}\|_\infty \leq m_v$ ▷ [from Equation (8)]
 calculate $\tilde{\sigma}_V^2 \leftarrow \sum_{\emptyset \neq v \in V \subseteq \{1, \dots, N\}} \sum_{\mathbf{j}_{|v|} \in \mathbb{N}_0^{|v|}, \|\mathbf{j}_{|v|}\|_\infty \leq m_v} C_{v,j_v}^2$ ▷ [from Equation (32)]
 calculate $\tilde{G}_{v,m_v} \leftarrow \left(\sum_{\mathbf{j}_{|v|} \in \mathbb{N}_0^{|v|}, \|\mathbf{j}_{|v|}\|_\infty \leq m_v} C_{v,j_v}^2 \right) / \tilde{\sigma}_V^2$ ▷ [from Equation (24)]
 rank $y_v(\mathbf{x}_v)$: $y_{v^{(1)}}(\mathbf{x}_{v^{(1)}})$ to $y_{v^{(N)}}(\mathbf{x}_{v^{(N)}})$ ▷ [from Algorithm 2]
 Get L ▷ [from Algorithm 2]
 $N_{m_u} \leftarrow 0$
 for $i \leftarrow 1$ to L **do** ▷ [comparing rankings from m_u with those from $(m_u - 1)$ to check for convergence]
 $R_{m_u}(i) \leftarrow i$
 if $R_{m_u-1}(i) = R_{m_u}(i)$ **then** $N_{m_u} \leftarrow N_{m_u} + 1$
 end if
 end for
 until $N_m/L \geq \epsilon_3$ ▷ [ranking converge]
 for $l_u \leftarrow 1$ to L **do** ▷ [start the adaptivity algorithm with the highest ranking $|u|$ -variate component function]
 $u \leftarrow u^{(l_u)}$
 repeat ▷ [continue incrementing the polynomial order m_u unless the adaptivity condition $\Delta \tilde{G}_{u,m_u} > \epsilon_2$ fails]
 $m_u \leftarrow m_u + 1$
 calculate $C_{u,j_u}, \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, \|\mathbf{j}_{|u|}\|_\infty \leq m_u$ ▷ [from Equation (8)]
 calculate $\tilde{\sigma}_V^2 \leftarrow \sum_{\emptyset \neq v \in V \subseteq \{1, \dots, N\}} \sum_{\mathbf{j}_{|v|} \in \mathbb{N}_0^{|v|}, \|\mathbf{j}_{|v|}\|_\infty \leq m_v} C_{v,j_v}^2$ ▷ [from Equation (32)]
 calculate $\tilde{G}_{u,m_u} \leftarrow \left(\sum_{\mathbf{j}_{|v|} \in \mathbb{N}_0^{|v|}, \|\mathbf{j}_{|v|}\|_\infty \leq m_v} C_{v,j_v}^2 \right) / \tilde{\sigma}_V^2$ ▷ [from Equation (24)]
 calculate $\Delta \tilde{G}_{u,m_u} \leftarrow (\tilde{G}_{u,m_u} - \tilde{G}_{u,m_u-1}) / \tilde{G}_{u,m_u-1}$ ▷ [from Equation (25)]
 until $\Delta \tilde{G}_{u,m_u} \leq \epsilon_2$
 if $\tilde{G}_{u,m_u} \leq \epsilon_1$ **then** **exit**
 end if ▷ [exit the adaptivity algorithm]
 end for
 end for
 calculate y_\emptyset ▷ [from Equation (2)]

replacing the exact variance σ^2 in Equations (24) and (25) facilitates an effective iterative scheme for estimating \tilde{G}_{u,m_u} and $\Delta \tilde{G}_{u,m_u}$ as well. Equation (32) was implemented in the proposed algorithm, as explained in Algorithm 1 and Figure 1.

The second stage of the algorithm requires an efficient procedure for selecting appropriate PDD component functions that are retained in an adaptive-sparse PDD approximation. For a given $1 \leq |u| \leq N$, let $y_u(\mathbf{X}_u)$, $\emptyset \neq u \subseteq \{1, \dots, N\}$ denote all $|u|$ -variate non-constant PDD component functions of y . It is elementary to count the number of these component functions to be $L_{|u|} = \binom{N}{|u|}$. Depending on the tolerance criteria specified, some or none of these component functions may contribute towards the resultant PDD approximation. Since the component functions are not necessarily hierarchically arranged, determining their relative significance to PDD approximation is not straightforward. Therefore, additional efforts to rank the

component functions are needed, keeping in mind that the same efforts may be recycled for the PDD approximation. For this purpose, the authors propose two distinct ranking schemes: (1) full ranking scheme and (2) a reduced ranking scheme, both exploiting the global sensitivity index G_u as a measure of the significance of $y_u(\mathbf{X}_u)$. However, since G_u is estimated by its m_u -th-order polynomial approximation \tilde{G}_{u,m_u} , any ranking system based on \tilde{G}_{u,m_u} , where m_u is finite, may be in a flux and should hence be carefully interpreted. This implies that a ranking scheme resulting from \tilde{G}_{u,m_u} , whether full or reduced, must be iterated for increasing values of m_u until the ranking scheme converges according to a specified criterion. In the full ranking scheme, all $|u|$ -variate component functions are re-ranked from scratch for each increment of m_u until a converged ranking scheme emerges. Consequently, the full ranking scheme affords any component function to contribute to the resultant

PDD approximation, provided that the criterion $\tilde{G}_{u,m_u} > \epsilon_1$ is satisfied only at convergence. In contrast, a subset of $|u|$ -variate component functions, determined from the previous ranking results and truncations set by the tolerance criterion, are re-ranked for each increment of m_u in the reduced ranking scheme until convergence is achieved. Therefore, for a component function from the reduced ranking scheme to contribute to the resultant PDD approximation, the criterion $\tilde{G}_{u,m_u} > \epsilon_1$ must be satisfied at all ranking iterations including the converged one. Therefore, the full ranking scheme is meticulous, but it is also exhaustive, rapidly becoming inefficient or impractical when applied to high-dimensional stochastic responses. The reduced ranking scheme, obtained less rigorously than the former, is highly efficient and is ideal for solving industrial-scale high-dimensional problems. A ranking system obtained at $m_u = m$, $2 \leq m < \infty$, for all $|u|$ -variate component functions is considered to be converged if the ranking discrepancy ratio, defined as the ratio of the number of ranked positions changed when m_u increases from $m - 1$ to m to the number of component functions ranked at $m_u = m - 1$, does not exceed the ranking tolerance $0 \leq \epsilon_3 \leq 1$. The number of component functions ranked in the full ranking scheme is $L_{|u|}$, the total number of $|u|$ -variate component functions, and is the same for any m_u or function y . In contrast, the number of component functions ranked in the reduced ranking scheme, which is equal to or less than $L_{|u|}$, depends on m_u , y , and ϵ_1 . Both ranking schemes are described in Algorithm 2.

3.5.2. Computational Effort

For uncertainty quantification, the computational effort is commonly determined by the total number of original function evaluations. Consequently, the efforts required by the proposed methods are proportional to the total numbers of the PDD expansion coefficients retained in the concomitant approximations and depend on the numerical techniques used to calculate the coefficients. The numerical evaluation of the expansion coefficients are discussed in Section 4.

The numbers of coefficients by the fully and partially adaptive-sparse PDD methods are

$$\begin{aligned} \bar{K} &= 1 + \sum_{\emptyset \neq u \subseteq \{1, \dots, N\}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_{|u|}\|_{\infty}=m_u, j_1, \dots, j_{|u|} \neq 0 \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2}} 1 \\ &= 1 + \sum_{\emptyset \neq u \subseteq \{1, \dots, N\}} \sum_{m_u=1}^{\infty} \sum_{\tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2} \left[m_u^{|u|} - (m_u - 1)^{|u|} \right] \end{aligned} \quad (33)$$

and

$$\begin{aligned} \bar{K}_S &= 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_{|u|}\|_{\infty}=m_u, j_1, \dots, j_{|u|} \neq 0 \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2}} 1 \\ &= 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{m_u=1}^{\infty} \sum_{\tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2} \left[m_u^{|u|} - (m_u - 1)^{|u|} \right], \end{aligned}$$

respectively. It is elementary to show that $\bar{K}_S \leq \bar{K}$ when $S \leq N$ for identical tolerances, as expected, with equality when $S = N$. Therefore, a partially adaptive-sparse PDD method in general is more economical than the fully adaptive-sparse PDD method.

What can be inferred from the numbers of coefficients required by a partially adaptive-sparse PDD method and the existing truncated PDD method? The following two results, Proposition 2 and 3, provide some insights when the tolerances vanish and when the largest orders of polynomials are identical.

Proposition 2. *If $\epsilon_1 \rightarrow 0$, and $\epsilon_2 \rightarrow 0$, then $\bar{K}_S \rightarrow \tilde{K}_{S,m}$ as $m \rightarrow \infty$.*

Proof. From Equation (34),

$$\begin{aligned} \lim_{\substack{\epsilon_1 \rightarrow 0 \\ \epsilon_2 \rightarrow 0}} \bar{K}_S &= 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_{|u|}\|_{\infty}=m_u \\ j_1, \dots, j_{|u|} \neq 0}} 1 \\ &= 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|} \\ j_1, \dots, j_{|u|} \neq 0}} 1 \\ &= \lim_{m \rightarrow \infty} \left[1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, \|\mathbf{j}_{|u|}\|_{\infty} \leq m \\ j_1, \dots, j_{|u|} \neq 0}} 1 \right] \quad (35) \\ &= \lim_{m \rightarrow \infty} \left[\sum_{k=0}^S \binom{N}{k} m^k \right] \\ &= \lim_{m \rightarrow \infty} \tilde{K}_{S,m}, \end{aligned}$$

where the last line follows from Equation (19). \square

Proposition 3. *If*

$$m_{\max} = \max_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\}, 1 \leq |u| \leq S \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2}} m_u < \infty \quad (36)$$

is the largest order of polynomial expansion for any component function $y_u(\mathbf{X}_u)$, $\emptyset \neq u \subseteq \{1, \dots, N\}$, $1 \leq |u| \leq S$, such that $\tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2$, then $\bar{K}_S \leq \tilde{K}_{S,m_{\max}}$.

Proof. From Equation (34),

$$\begin{aligned} \bar{K}_S &= 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{m_u=1}^{\infty} \sum_{\substack{\|\mathbf{j}_{|u|}\|_{\infty}=m_u, j_1, \dots, j_{|u|} \neq 0 \\ \tilde{G}_{u,m_u} > \epsilon_1, \Delta \tilde{G}_{u,m_u} > \epsilon_2}} 1 \\ &\leq 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ 1 \leq |u| \leq S}} \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, \|\mathbf{j}_{|u|}\|_{\infty} \leq m_{\max} \\ j_1, \dots, j_{|u|} \neq 0}} 1 \quad (37) \\ &= \sum_{k=0}^S \binom{N}{k} m_{\max}^k \\ &= \tilde{K}_{S,m_{\max}}, \end{aligned}$$

where the last line follows from Equation (19). \square

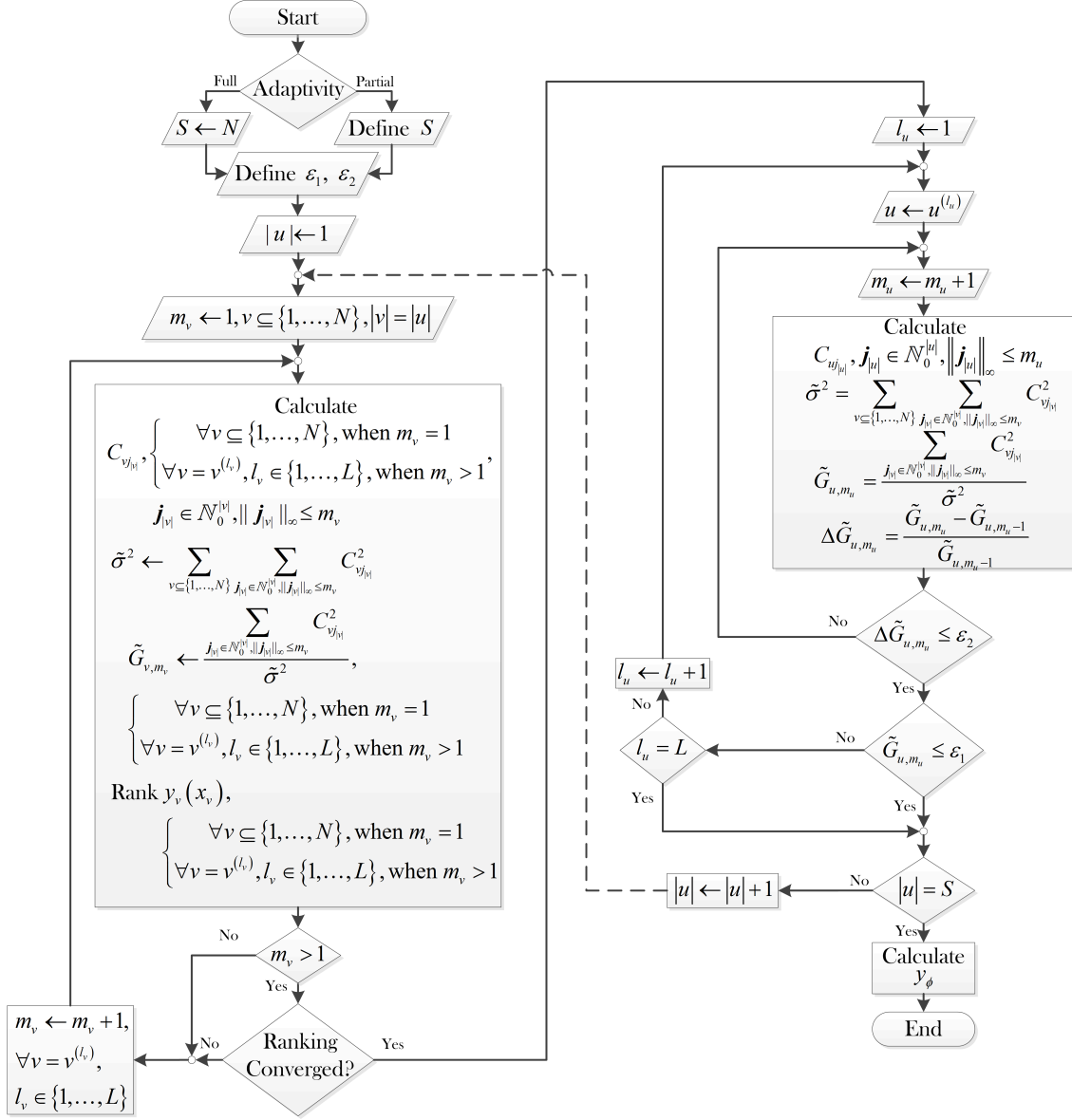


Figure 1: A flowchart for constructing an adaptive-sparse polynomial dimensional decomposition.

Algorithm 2 Ranking of component functions.

```

sort  $y_{v^{(l)}}(\mathbf{x}_{v^{(l)}})$ :  $l = 1, \dots, L$ ;  $l = 1$  for largest  $\tilde{G}_{v, m_v}$  ▷  $[L = N$  for full ranking, or when  $m_v = 1]$ 
Truncation for reduced ranking:
 $l \leftarrow 1$ 
while  $\tilde{G}_{v^{(l)}, m_v^{(l)}} > \epsilon_l$  do ▷ [truncating the ranking when adaptivity condition  $\tilde{G}_{v^{(l)}, m_v^{(l)}} > \epsilon_l$  fails]
     $L \leftarrow l$ 
     $l \leftarrow l + 1$ 
end while

```

According to Proposition 3, the partially adaptive-sparse PDD approximation for non-trivial tolerances should be computationally more efficient than the truncated PDD approximation. However, the computational efforts by both approximations depend on the numerical technique employed to estimate the associated expansion coefficients. For instance, suppose that a full-grid dimension-reduction integration with its own truncation $R = S$, to be explained in Section 4, is applied to calculate all $\tilde{K}_{S, m_{\max}}$ expansion coefficients to achieve the accuracy of an S -variate, m_{\max} -th-order PDD approximation. Then the requisite number of function evaluations is S -th-order polynomial with respect to N , the size of the stochastic problem. The partially adaptive-sparse PDD approximation, while retaining a similar accuracy, is expected to markedly reduce the number of function calls. This issue will be further explored in Example 3 of the Numerical Examples section.

4. Calculation of Expansion Coefficients

The determination of the expansion coefficients y_0 and $C_{i\mathbf{j}_{|u|}}$ in Equations (2) and (8) involves various N -dimensional integrals over \mathbb{R}^N . For large N , a full numerical integration employing an N -dimensional tensor product of a univariate quadrature formula is computationally prohibitive and is, therefore, ruled out. Two new alternative numerical techniques are proposed to estimate the coefficients accurately and efficiently.

4.1. Dimension-Reduction Integration

The dimension-reduction integration, developed by Xu and Rahman [40], entails approximating a high-dimensional integral of interest by a finite sum of lower-dimensional integrations. For calculating the expansion coefficients y_0 and $C_{i\mathbf{j}_{|u|}}$, this is accomplished by replacing the N -variate function y in Equations (2) and (8) with an R -variate RDD approximation at a chosen reference point, where $R \leq N$ [40, 41]. The result is a reduced integration scheme, requiring evaluations of at most R -dimensional integrals.

Given a reference point $\mathbf{c} = (c_1, \dots, c_N) \in \mathbb{R}^N$ and RDD component functions w_0 and $w_u(\mathbf{X}_u; \mathbf{c})$ described by Equations (5) and (6), let $\hat{y}_R(\mathbf{X}; \mathbf{c})$ (Equation (11)) denote an R -variate RDD approximation of $y(\mathbf{X})$. Replacing $y(\mathbf{x})$ in Equations (2) and (8) with $\hat{y}_R(\mathbf{x}; \mathbf{c})$, the coefficients y_0 and $C_{i\mathbf{j}_{|u|}}$ are estimated from [40]

$$y_0 \cong \sum_{i=0}^R (-1)^i \binom{N-R+i-1}{i} \sum_{\substack{v \subseteq \{1, \dots, N\} \\ |v|=R-i}} \int_{\mathbb{R}^{|v|}} y(\mathbf{x}_v, \mathbf{c}_{-v}) f_{\mathbf{X}_v}(\mathbf{x}_v) d\mathbf{x}_v$$

and

$$C_{i\mathbf{j}_{|u|}} \cong \sum_{i=0}^R (-1)^i \binom{N-R+i-1}{i} \sum_{\substack{v \subseteq \{1, \dots, N\} \\ |v|=R-i, u \subseteq v}} \int_{\mathbb{R}^{|v|}} y(\mathbf{x}_v, \mathbf{c}_{-v}) \psi_{i\mathbf{j}_{|u|}}(\mathbf{x}_u) f_{\mathbf{X}_v}(\mathbf{x}_v) d\mathbf{x}_v, \quad (39)$$

respectively, requiring evaluation of at most R -dimensional integrals. The reduced integration facilitates calculation of the coefficients approaching their exact values as $R \rightarrow N$, and is significantly more efficient than performing one N -dimensional integration, particularly when $R \ll N$. Hence, the computational effort is significantly decreased using the dimension-reduction integration. For instance, when $R = 1$ or 2 , Equations (38) and (39) involve one-, or at most, two-dimensional integrations, respectively. Nonetheless, numerical integrations are still required for performing various $|v|$ -dimensional integrals over $\mathbb{R}^{|v|}$, where $0 \leq |v| \leq R$. When $R > 1$, the multivariate integrations involved can be conducted using full- or sparse-grids, as follows.

4.1.1. Full-Grid Integration

The full-grid dimension-reduction integration entails constructing a tensor product of the underlying univariate quadrature rules. For a given $v \subseteq \{1, \dots, N\}$, $1 < |v| \leq R$, let $v = \{i_1, \dots, i_{|v|}\}$, where $1 \leq i_1 < \dots < i_{|v|} \leq N$. Denote by $\{x_{i_p}^{(1)}, \dots, x_{i_p}^{(n_v)}\} \subset \mathbb{R}$ a set of integration points of x_{i_p} and by $\{w_{i_p}^{(1)}, \dots, w_{i_p}^{(n_v)}\}$ the associated weights generated from a chosen univariate quadrature rule and a positive integer $n_v \in \mathbb{N}$. Denote by $\mathbf{P}^{(n_v)} = \times_{p=1}^{|v|} \{x_{i_p}^{(1)}, \dots, x_{i_p}^{(n_v)}\}$ the rectangular grid consisting of all integration points generated by the variables indexed by the elements of v . Then the coefficients using dimension-reduction numerical integration with a full-grid are approximated by

$$y_0 \cong \sum_{i=0}^R (-1)^i \binom{N-R+i-1}{i} \sum_{\substack{v \subseteq \{1, \dots, N\} \\ |v|=R-i}} \sum_{\mathbf{k}_{|v|} \in \mathbf{P}^{(n_v)}} w^{(\mathbf{k}_{|v|})} y(\mathbf{x}_v^{(\mathbf{k}_{|v|})}, \mathbf{c}_{-v}), \quad (40)$$

$$C_{u\mathbf{j}_{|v|}} \cong \sum_{i=0}^R (-1)^i \binom{N-R+i-1}{i} \sum_{\substack{v \subseteq \{1, \dots, N\} \\ |v|=R-i, u \subseteq v}} \sum_{\mathbf{k}_{|v|} \in P^{(n_v)}} w^{(\mathbf{k}_{|v|})} y(\mathbf{x}_v^{(\mathbf{k}_{|v|})}, \mathbf{c}_{-v}) \psi_{u\mathbf{j}_{|v|}}(\mathbf{x}_u^{(\mathbf{k}_{|v|})}), \quad (41)$$

where $\mathbf{x}_v^{(\mathbf{k}_{|v|})} = \{x_{i_1}^{(k_{i_1})}, \dots, x_{i_{|v|}}^{(k_{i_{|v|}})}\}$ and $w^{(\mathbf{k}_{|v|})} = \prod_{p=1}^{|v|} w_{i_p}^{(k_p)}$ is the product of integration weights generated by the variables indexed by the elements of v . For independent coordinates of \mathbf{X} , as assumed here, a univariate Gauss quadrature rule is commonly used, where the integration points and associated weights depend on the probability distribution of X_i . They are readily available, for example, the Gauss-Hermite or Gauss-Legendre quadrature rule, when X_i follows Gaussian or uniform distribution [11]. For an arbitrary probability distribution of X_i , the Stieltjes procedure [11] can be employed to generate the measure-consistent Gauss quadrature formulae [11]. An n_v -point Gauss quadrature rule exactly integrates a polynomial of total degree at most $2n_v - 1$.

The calculation of y_0 and $C_{u\mathbf{j}_{|v|}}$ from Equations (40) and (41) involves at most R -dimensional tensor products of an n_v -point univariate quadrature rule, requiring the following deterministic responses or function evaluations: $y(\mathbf{c})$, $y(\mathbf{x}_v^{(\mathbf{j}_{|v|})}, \mathbf{c}_{-v})$ for $i = 0, \dots, R$, $v \subseteq \{1, \dots, N\}$, $|v| = R - i$, and $\mathbf{j}_{|v|} \in P^{(n_v)}$. Accordingly, the total cost for estimating the PDD expansion coefficients entails

$$L_{FG} = \sum_{i=0}^R \sum_{\substack{v \subseteq \{1, \dots, N\} \\ |v|=R-i}} n_v^{|v|} \quad (42)$$

function evaluations, encountering a computational complexity that is R th-order polynomial – for instance, linear or quadratic when $R = 1$ or 2 – with respect to the number of random variables or integration points. For $R < N$, the technique alleviates the curse of dimensionality to an extent determined by R .

4.1.2. Sparse-Grid Integration

Although the full-grid dimension-reduction integration has been successfully applied to the calculation of the PDD expansion coefficients in the past [25–27, 32], it faces a major drawback when the polynomial order m_u for a PDD component function y_u needs to be modulated for adaptivity. As the value of m_u is incremented by one, a completely new set of integration points is generated by the univariate Gauss quadrature rule, rendering all expensive function evaluations on prior integration points as useless. Therefore, a nested Gauss quadrature rule, such as the fully symmetric interpolatory rule, that is capable of exploiting dimension-reduction integration is proposed.

Fully symmetric interpolatory rule. The fully symmetric interpolatory (FSI) rules developed by Genz and his associates [12, 13], is a sparse-grid integration technique for performing high-dimensional numerical integration. Applying this rule to

the $|v|$ -dimensional integrations in Equations (38) and (39), the PDD expansion coefficients are approximated by

$$y_0 \cong \sum_{i=0}^R (-1)^i \binom{N-R+i-1}{i} \sum_{\substack{v \subseteq \{1, \dots, N\} \\ |v|=R-i}} \sum_{\mathbf{p}_{|v|} \in P^{(\tilde{n}_v, |v|)}} w_{\mathbf{p}_{|v|}} \quad (43)$$

$$\sum_{\mathbf{q}_{|v|} \in \Pi_{\mathbf{p}_{|v|}}} \sum_{\mathbf{t}_{|v|}} y(t_{i_1} \alpha_{q_{i_1}}, \dots, t_{i_{|v|}} \alpha_{q_{i_{|v|}}}, \mathbf{c}_{-v}),$$

$$C_{u\mathbf{j}_{|v|}} \cong \sum_{i=0}^R (-1)^i \binom{N-R+i-1}{i} \sum_{\substack{v \subseteq \{1, \dots, N\} \\ |v|=R-i, u \subseteq v}} \sum_{\mathbf{p}_{|v|} \in P^{(\tilde{n}_v, |v|)}} w_{\mathbf{p}_{|v|}} \sum_{\mathbf{q}_{|v|} \in \Pi_{\mathbf{p}_{|v|}}} \sum_{\mathbf{t}_{|v|}} y(t_{i_1} \alpha_{q_{i_1}}, \dots, t_{i_{|v|}} \alpha_{q_{i_{|v|}}}, \mathbf{c}_{-v}) \psi_{u\mathbf{j}_{|v|}}(t_{i_1} \alpha_{q_{i_1}}, \dots, t_{i_{|v|}} \alpha_{q_{i_{|v|}}}), \quad (44)$$

where $v = \{i_1, \dots, i_{|v|}\}$, $\mathbf{t}_{|v|} = (t_{i_1}, \dots, t_{i_{|v|}})$, $\mathbf{p}_{|v|} = (p_{i_1}, \dots, p_{i_{|v|}})$,

$$P^{(\tilde{n}_v, |v|)} = \{\mathbf{p}_{|v|} : \tilde{n}_v \geq p_{i_1} \geq \dots \geq p_{i_{|v|}} \geq 0, \|\mathbf{p}_{|v|}\| \leq \tilde{n}_v\} \quad (45)$$

with $\|\mathbf{p}_{|v|}\| := \sum_{r=1}^{|v|} p_{i_r}$ is the set of all distinct $|v|$ -partitions of the integers $0, 1, \dots, \tilde{n}_v$, and $\Pi_{\mathbf{p}_{|v|}}$ is the set of all permutations of $\mathbf{p}_{|v|}$. The innermost sum over $\mathbf{t}_{|v|}$ is taken over all of the sign combinations that occur when $t_{i_r} = \pm 1$ for those values of i_r with generators $\alpha_{q_{i_r}} \neq 0$ [13]. The weight

$$w_{\mathbf{p}_{|v|}} = 2^{-K} \sum_{\|\mathbf{k}_{|v|}\| \leq \tilde{n}_v - \|\mathbf{p}_{|v|}\|} \prod_{r=1}^{|v|} \frac{a_{k_{i_r} + p_{i_r}}}{\prod_{j=0, j \neq p_{i_r}} (a_{p_{i_r}}^2 - a_j^2)}, \quad (46)$$

where K is the number of nonzero components in $\mathbf{p}_{|v|}$ and a_i is a constant that depends on the probability measure of X_i , for instance,

$$a_i = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{\xi^2}{2}\right) \prod_{j=0}^{i-1} (\xi^2 - \alpha_j^2) d\xi \quad (47)$$

for $i > 0$ and $a_0 = 1$ when X_i follows the standard Gaussian distribution [13]. An \tilde{n}_v -parameter FSI rule exactly integrates a polynomial of degree at most $2\tilde{n}_v + 1$.

Extended fully symmetric interpolatory rule. The number of function evaluations by the original FSI rule [12] increases rapidly as $|v|$ and \tilde{n}_v increase. To enhance the efficiency, Genz and Keister [13] proposed an extended FSI rule in which the function evaluations are significantly reduced if the generator set is chosen such that some of the weights $w_{\mathbf{p}_{|v|}}$ are zero. The pivotal step in constructing such FSI rule is to extend a $(2\beta + 1)$ -point Gauss-Hermite quadrature rule by adding 2γ points or generators $\pm\alpha_{\beta+1}, \pm\alpha_{\beta+2}, \dots, \pm\alpha_{\beta+\gamma}$ with the objective of maximizing the degree of polynomial exactness of the extended rule, where $\beta \in \mathbb{N}$ and $\gamma \in \mathbb{N}$. Genz and Keister [13] presented a special case of initiating the FSI rule from the univariate Gauss-Hermite rule over the interval $(-\infty, \infty)$. The additional generators in this case are determined as roots of the monic polynomial

$\zeta^{2\gamma} + t_{\gamma-1}\zeta^{2\gamma-1} + \dots + t_0$, where the coefficients $t_{\gamma-1}, \dots, t_0$ are obtained by invoking the condition

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{\xi^2}{2}\right) \prod_{j=0}^{\beta} \xi^{2b} (\xi^2 - \alpha_j^2) d\xi = 0, \quad (48)$$

where $\gamma > \beta$. A new set of generators is propagated based on the prior rule and, therefore, as the polynomial degree of exactness of the rule increases, all the previous points and the expensive function evaluations over those points are preserved. A remarkable feature of the extended FSI rule is that the choice of generators is such that some of the weights $w_{\mathbf{p}_i} = 0$ in each step of the extension [13], thus eliminating the need for function evaluations at the integration points corresponding to *zero* weights, making the extended FSI rule significantly more efficient than its earlier version.

Since RDD is tied with the reference point, the dimension-reduction integration, whether full-grid or sparse-grid, to calculate the PDD expansion coefficients also depends on \mathbf{c} . However, from past experience [20, 25, 26, 40, 43], very accurate estimates of the expansion coefficients were obtained when \mathbf{c} is selected as the mean value of \mathbf{X} . A more rigorous approach entails finding an optimal reference point, but it will require additional function evaluations and hence may render the dimension-reduction technique impractical for solving high-dimensional problems.

4.1.3. Integration Points

The number of integration points determines the computational expense incurred in calculating the PDD expansion coefficients. Therefore, it is instructive to compare the numbers of points required by full- or sparse-grid dimension-reduction integrations. To do so, consider the efforts in performing a $|\mathbf{v}|$ -dimensional integration in Equation (38) or (39) over the interval $(-\infty, \infty)$ by three different numerical techniques: (1) the full-grid integration technique; (2) the sparse-grid integration technique using the extended FSI rule; and (3) the sparse-grid integration technique using Smolyak's algorithm [22]. The Smolyak's algorithm is included because it is commonly used as a preferred sparse-grid numerical technique for approximating high-dimensional integrals. Define an integer $l \in \mathbb{N}$ such that all three techniques can exactly integrate a polynomial function of total degree $2l - 1$. For instance, when $l = 3$, all three techniques exactly integrate a quintic polynomial. Figure 2 presents a comparison of the total numbers of integration points in a two-dimensional grid, that is, when $|\mathbf{v}| = 2$, for l ranging from one through five by the three distinct multivariate integration techniques. Each plot illustrates two numbers: the first number indicates the number of integration points required at the given value of l ; the second number, inside the parenthesis, indicates the total number of cumulative integration points added up to the value of l . It is imperative to add the integration points from all the previous values of l as it reflects the total number of function evaluations required in an adaptive algorithm. For the full-grid integration, the two numbers are different for all $l > 1$, indicating a lack of nesting of the integration points. Whereas in the sparse-grid with extended FSI rule,

the two numbers are equal for all l , reflecting the fully nested integration points in this rule. As l increments, a completely new set of points is introduced in the full-grid integration, rendering the prior points useless. However, for fairness in comparison, it is necessary to consider all points from prior values of l as the expensive function evaluations have already been performed. Therefore, Figure 2 captures the cumulative numbers of integration points as l increases steadily. For values of l up to two, all three techniques require the same number of integration points. However, differences in the numbers of points start to appear in favor of the extended FSI rule when l exceeds two, making it the clear favorite among all three techniques for high-order numerical integration. The Smolyak's algorithm, which is not nested, is the least efficient of the three techniques. The extended FSI rule, in contrast, is fully nested, establishing a principal advantage over Smolyak's algorithm for adaptive numerical integration.

Table 1 lists the number of integration points required at the integration rule corresponding to a given value of l , for $2 \leq |\mathbf{v}| \leq 10$ and $2 \leq l \leq 5$. It is important to note that the number of integration points listed is not cumulative. It appears that for higher-dimensional integrations, that is, for $|\mathbf{v}| > 2$, the extended FSI rule is markedly more efficient than full-grid or other sparse-grid techniques even for the non-cumulative points. The efficiency of extended FSI rule is more pronounced for cumulative number of integration points. For further details, the reader is referred to the work of Genz and Keister [13], who examined the extended FSI rule for dimensions up to 20.

4.2. Quasi Monte Carlo Simulation

The basic idea of the quasi MCS is to replace the random or pseudo-random samples in crude MCS by well-chosen deterministic samples that are highly equidistributed [21]. The quasi Monte Carlo samples are often selected from a low-discrepancy sequence [9, 16, 21, 34] or by a lattice rule [33] to minimize the integration errors. The estimation of the PDD expansion coefficients, which are high-dimensional integrals, comprises three simple steps: (1) generate a low-discrepancy point set $\mathcal{P}_L := \{\mathbf{u}^{(k)} \in [0, 1]^N, k = 1, \dots, L\}$ of size $L \in \mathbb{N}$; (2) map each sample from \mathcal{P}_L to the sample $\mathbf{x}^{(k)} \in \mathbb{R}^N$ following the probability measure of the random input \mathbf{X} ; and (3) approximate the coefficients by

$$y_0 \cong \frac{1}{L} \sum_{k=1}^L y(\mathbf{x}^{(k)}), \quad (49)$$

$$C_{u_{|\mathbf{j}|}} \cong \frac{1}{L} \sum_{k=1}^L y(\mathbf{x}^{(k)}) \psi_{u_{|\mathbf{j}|}}(\mathbf{x}_u^{(k)}). \quad (50)$$

The well-known Koksma–Hlawka inequality reveals that the error committed by the quasi MCS is bounded by the variation of the integrand in the sense of Hardy and Krause and the star-discrepancy, a measure of uniformity, of the point set \mathcal{P}_L [21]. Therefore, constructing a point set with star-discrepancy as small as possible and seeking variance reduction of the integrand are vital for the success of the quasi MCS. It should be

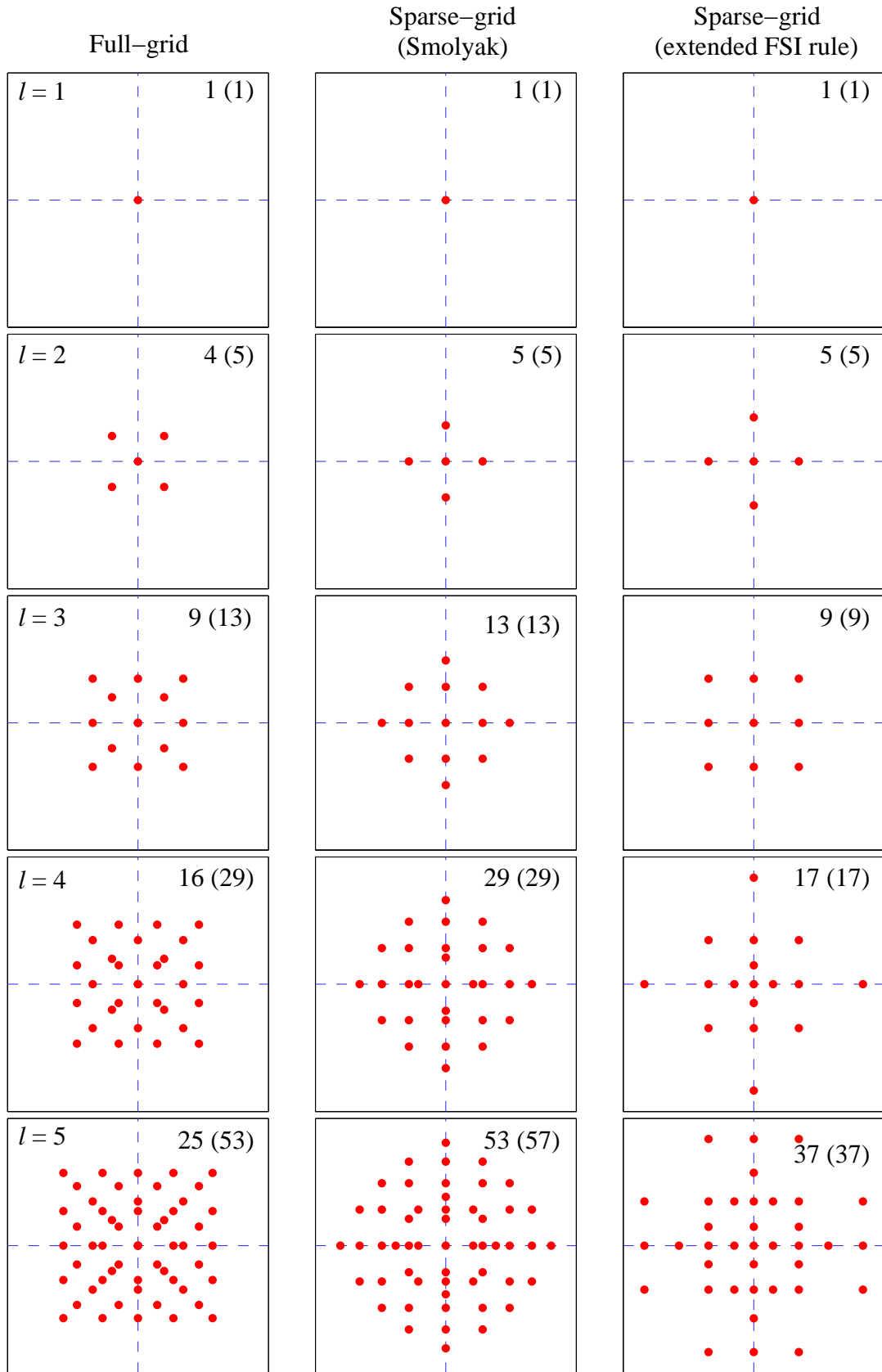


Figure 2: Gauss-Hermite integration points in a two-dimensional grid by the full-grid technique, sparse-grid with the extended FSI rule, and sparse-grid with Smolyak's algorithm for various levels. Note: each grid is plotted over a square with axes ranging from -5 to 5 .

Table 1: Number of integration points in various $|v|$ -dimensional integration techniques, each technique exactly integrates polynomials of total order $2l - 1$.

l	$ v $									
	2	3	4	5	6	7	8	9	10	
(a) Full-grid										
2	4	8	16	32	64	128	256	512	1024	
3	9	27	81	243	729	2187	6561	19683	59049	
4	16	64	256	1024	4096	16384	65536	262144	1048576	
5	25	125	625	3125	15625	78125	390625	1953125	9765625	
(b) Sparse-grid (Smolyak)										
2	5	7	9	11	13	15	17	19	21	
3	13	25	41	61	85	113	145	181	221	
4	29	69	137	241	389	589	849	1177	1581	
5	53	165	385	781	1433	2437	3905	5965	8761	
(c) Sparse-grid (extended FSI rule)										
2	5	7	9	11	13	15	17	19	21	
3	9	19	33	51	73	99	129	163	201	
4	17	39	81	151	257	407	609	871	1201	
5	37	93	201	401	749	1317	2193	3481	5301	

mentioned here that many authors, including Halton [16], Faure [9], Niederreiter [21], and Sobol [34], and Wang [38], have extensively studied how to generate the best low-discrepancy point sets and to facilitate variance reduction. For a bounded variation of the integrand, the quasi MCS has a theoretical error bound $O(L^{-1}(\log L)^N)$ compared with the probabilistic error bound $O(L^{-1/2})$ of crude MCS, indicating significantly faster convergence of the quasi MCS than crude MCS.

The two proposed techniques for calculating the PDD coefficients represent two broad categories of numerical integration: the quadrature-based methods and the sampling-based methods. However, the calculation of PDD coefficients is not limited to only these two techniques. Furthermore, the relative accuracy or efficiency of one technique over the other depends on the dimension of the stochastic problem. For hundreds or thousands of random variables, a sampling-based technique is generally preferred over a quadrature-based technique, as the former is relatively insensitive to the problem size.

5. Numerical Examples

Three numerical examples are put forward to illustrate the adaptive-sparse PDD methods developed in calculating various probabilistic characteristics of random mathematical functions and random eigensolutions of stochastic dynamical systems. A principal objective is to compare the performance of the proposed adaptive-sparse PDD methods with that of the existing truncated PDD method. Readers interested in contrasting the truncated PDD method with the PCE [14] and other classical methods are referred to the authors' prior work [24–26, 32].

Classical Legendre polynomials were used to define the orthonormal polynomials in Example 1, and all expansion coefficients were determined analytically. In Examples 2 and 3, all original random variables were transformed into standard Gaussian random variables, facilitating the use of classical Hermite orthonormal polynomials as bases. Since Example 2 con-

sists of only nine input random variables, the expansion coefficients were estimated using a nine-dimensional tensor product of five-point univariate Gauss-Hermite quadrature rule. The expansion coefficients in Example 3 were approximated by both the full-grid dimension-reduction integration and sparse-grid dimension-reduction integration with the extended FSI rule, where $R = S$ and \mathbf{c} is the mean of \mathbf{X} . The sample sizes for crude MCS in Example 2 is 10^6 . In Example 3, the sample size for crude MCS is 50,000, and for the embedded MCS, whether the truncated or adaptive-sparse PDD method, the sample size is 10^6 .

5.1. Example 1: A Polynomial Function

Consider the polynomial function

$$y(\mathbf{X}) = \frac{\prod_{i=1}^N \left(\frac{3}{i} X_i^5 + 1 \right)}{\mathbb{E} \left[\prod_{i=1}^N \left(\frac{3}{i} X_i^5 + 1 \right) \right]}, \quad (51)$$

where X_i , $i = 1, \dots, N$, are independent and identical random variables, each following the standard uniform distribution over $[0, 1]$. Since the coefficient of X_i^5 is inversely proportional to i , the first and last random variables have the largest and least influence on y . From elementary calculations, the exact mean and variance of y are 1 and

$$\prod_{i=1}^N \left(\frac{25}{11(1+2i)^2} + 1 \right) - 1, \quad (52)$$

respectively. All PDD expansion coefficients were calculated analytically. Therefore, the ranking of component functions was performed once and for all, avoiding any role of the ranking scheme in this particular example. The numerical results that follow in the remainder of this subsection were obtained for $N = 5$.

Figure 3 shows how the relative errors, defined as the ratio of the absolute difference between the exact (Expression (52)) and approximate (Equation (17)) variances of y to the exact variance, committed by S -variate, m -th order PDD approximations vary with increasing polynomial order m . The five plots of univariate ($S = 1$) to pentavariate ($S = 5$) PDD approximations clearly show that the error drops monotonically with respect to m regardless of S . When m reaches five, the pentavariate PDD approximation does not perpetrate any error, producing the exact variance of y as expected. In contrast, the relative errors in variance caused by fully adaptive-sparse PDD approximations (Equation (30)), also illustrated in Figure 3 for specified tolerances ranging from 10^{-9} to 10^{-3} , do not rely on S or m , as the degrees of interaction and polynomial orders are adaptively modulated in the concomitant approximations. The adaptive-sparse PDD approximations with tolerances equal to 10^{-3} and 10^{-4} yield relative errors in variance marginally higher than the tolerance values; however, the relative errors achieved are invariably smaller than all respective values of the subsequent tolerances, demonstrating a one-to-one relationship between the tolerance and relative error attained in calculating the variance. As the tolerance decreases, so does the relative error. While a traditional truncated PDD approximation provides options to increase the values of S and/or m for reducing the relative error, the user remains blinded to the outcome of such an action. The adaptive-sparse PDD method, in the form of tolerances, provides a direct key to regulate the accuracy of the resultant approximation.

Figure 4 displays the increase in number of PDD expansion coefficients required by truncated (Equation (19)) and fully adaptive-sparse (Equation (33)) PDD methods in order to achieve a user-specified relative error in variance ranging from 10^{-1} to 10^{-12} . The relative error decreases from left to right along the horizontal axis of the plot. The plot of the truncated PDD approximation is generated by trial-and-error, increasing the value of either S or m until the desired relative error is achieved and then counting the total number of coefficients required to attain that relative error. For obtaining the plot of the adaptive-sparse PDD approximation, the tolerance values were reduced monotonically and the corresponding total number of coefficients was noted for each value of relative error. Ignoring the two lowest relative errors, the comparison of the plots from these two methods clearly demonstrates how the adaptive-sparse PDD method requires fewer expansion coefficients than the truncated PDD method to achieve the desired level of relative error. While the adaptive-sparse PDD method intelligently calculates only those coefficients that are making significant contribution to the variance, the truncated PDD method ends up calculating more coefficients than required. Therefore, the adaptive-sparse PDD approximation represents a more scientific and efficient method than the truncated PDD methods.

5.2. Example 2: Eigenvalues of an Undamped, Spring-Mass System

Consider a three-degree-of-freedom, undamped, spring-mass system, shown in Figure 5, with random mass and random stiff-

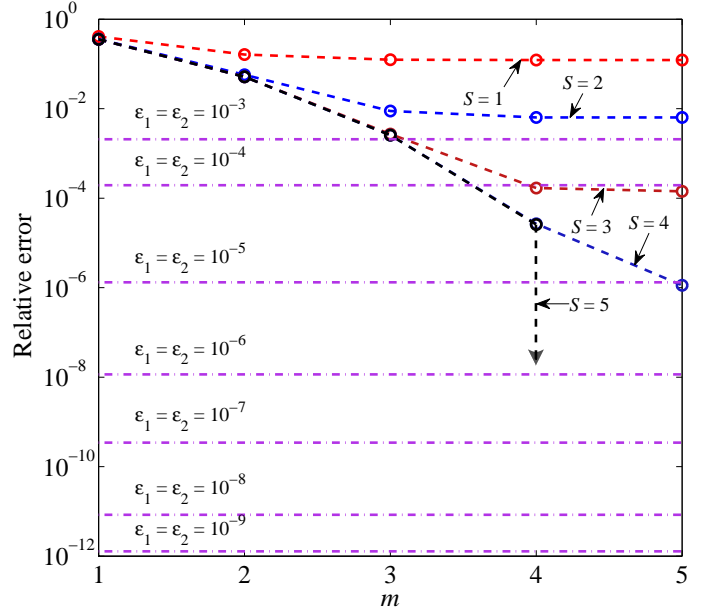


Figure 3: Relative error in calculating the variance of a mathematical function by fully adaptive-sparse and truncated PDD methods (Example 1).

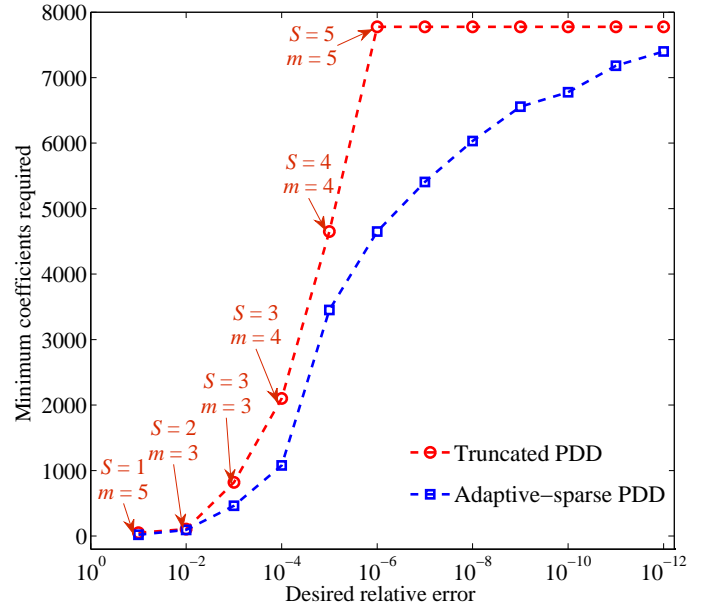


Figure 4: Minimum number of coefficients required to achieve a desired relative error in the variance of a mathematical function by fully adaptive-sparse and truncated PDD methods (Example 1).

ness matrices

$$\mathbf{M}(\mathbf{X}) = \begin{bmatrix} M_1(\mathbf{X}) & 0 & 0 \\ 0 & M_2(\mathbf{X}) & 0 \\ 0 & 0 & M_3(\mathbf{X}) \end{bmatrix} \quad (53)$$

and

$$\mathbf{K}(\mathbf{X}) = \begin{bmatrix} K_{11}(\mathbf{X}) & K_{12}(\mathbf{X}) & K_{13}(\mathbf{X}) \\ & K_{22}(\mathbf{X}) & K_{23}(\mathbf{X}) \\ (\text{sym.}) & & K_{33}(\mathbf{X}) \end{bmatrix}, \quad (54)$$

respectively, where $K_{11}(\mathbf{X}) = K_1(\mathbf{X}) + K_4(\mathbf{X}) + K_6(\mathbf{X})$, $K_{12}(\mathbf{X}) = -K_4(\mathbf{X})$, $K_{13}(\mathbf{X}) = -K_6(\mathbf{X})$, $K_{22}(\mathbf{X}) = K_4(\mathbf{X}) + K_5(\mathbf{X}) + K_2(\mathbf{X})$, $K_{23}(\mathbf{X}) = -K_5(\mathbf{X})$, and $K_{33}(\mathbf{X}) = K_5(\mathbf{X}) + K_3(\mathbf{X}) + K_6(\mathbf{X})$; the masses $M_i(\mathbf{X}) = \mu_i X_i$; $i = 1, 2, 3$ with $\mu_i = 1.0$ kg; $i = 1, 2, 3$, and spring stiffnesses $K_i(\mathbf{X}) = \mu_{i+3} X_{i+3}$; $i = 1, \dots, 6$ with $\mu_{i+3} = 1.0$ N/m; $i = 1, \dots, 5$ and $\mu_9 = 3.0$ N/m. The input $\mathbf{X} = \{X_1, \dots, X_9\}^T \in \mathbb{R}^9$ is an independent log-normal random vector with mean $\boldsymbol{\mu}_{\mathbf{X}} = \mathbf{1} \in \mathbb{R}^9$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{X}} = \nu^2 \mathbf{I} \in \mathbb{R}^{9 \times 9}$ with coefficient of variation $\nu = 0.3$.

Three partially adaptive-sparse PDD methods with $S = 1, 2$, and 3 were applied to calculate the variances (Equation (31)) of the three random eigenvalues of the dynamic system. The tolerance values are as follows: $\epsilon_1 = \epsilon_2 = 10^{-6}$ and $\epsilon_3 = 0.7$. Table 2 presents the variances of eigenvalues from various partially adaptive-sparse PDD methods calculated according to Algorithms 1 and 2. The results of both full and reduced ranking systems are tabulated. Also included in Table 2 are the variance calculations from crude MCS. The variances obtained using the univariate ($S = 1$) partially adaptive-sparse PDD approximation are relatively far from the benchmark results of crude MCS since the univariate approximation is unable to capture any interactive effects of the input variables. However, the bivariate ($S = 2$) and trivariate ($S = 3$) partially adaptive-sparse PDD approximations achieve very high accuracy in calculating the variances of all three random eigenvalues. Remarkably, the reduced ranking scheme delivers the same level of accuracy, at least up to three decimal places shown, of the full ranking scheme in calculating the variances.

In order to study the efficiency of the reduced ranking scheme vis-a-vis the full ranking scheme in a trivariate partially adaptive-sparse PDD approximation, the corresponding total numbers of coefficients (Equation (34)) required were compared, along with the total number of coefficients (Equation (19)) required in a trivariate, fifth-order truncated PDD approximation, in Figure 6. The order of the truncated PDD is the largest value of m_u required in the adaptive-sparse PDD approximation. While the partially adaptive-sparse PDD method with either ranking scheme requires fewer coefficients than does the truncated PDD method, it is the reduced ranking scheme that is the clear winner in efficiency with the least number of coefficients. The largest reduction in the number of coefficients achieved by the reduced ranking system is approximately sixty-eight percent when calculating the variance of the third eigenvalue. These results are in agreement with Proposition 3.

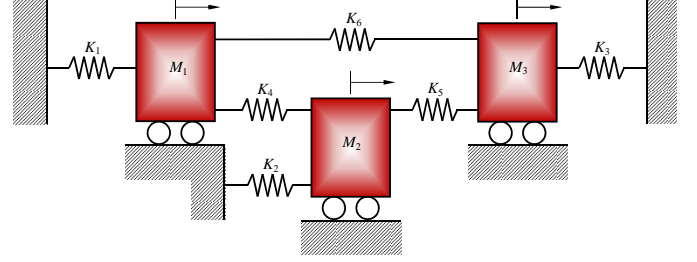


Figure 5: A three-degree-of-freedom undamped, spring-mass system (Example 2).

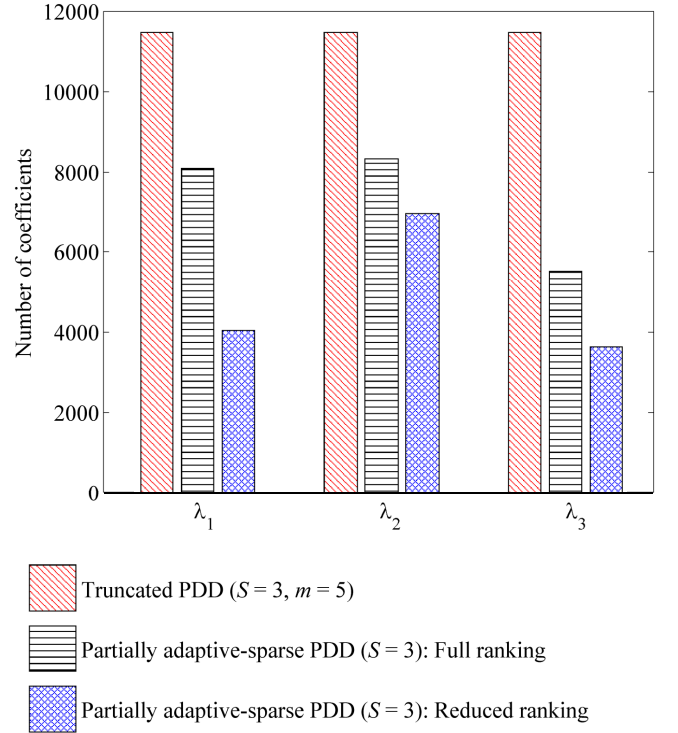


Figure 6: Number of coefficients required for calculating the variance of a three-degree-of-freedom linear oscillator by trivariate partially adaptive-sparse PDD approximations using full and reduced ranking schemes.

Table 2: Variances of three eigenvalues of a three-degree-of-freedom linear oscillator by three partially adaptive-sparse PDD methods and crude MCS.

	$S = 1$		$S = 2$		$S = 3$		MCS
λ	Full ranking	Reduced ranking	Full ranking	Reduced ranking	Full ranking	Reduced ranking	10^6
1	0.057	0.057	0.060	0.060	0.060	0.060	0.060
2	1.152	1.152	1.204	1.204	1.215	1.215	1.219
3	7.289	7.289	7.576	7.576	7.585	7.585	7.585

5.3. Example 3: Modal Analysis of a Functionally Graded Cantilever Plate

The third example involves free vibration analysis of a $2\text{ m} \times 1\text{ m} \times 10\text{ mm}$ cantilever plate, shown in Figure 7(a), made of a functionally graded material (FGM)², where silicon carbide (SiC) particles varying along the horizontal coordinate ξ are randomly dispersed in an aluminum (Al) matrix [42]. The result is a random inhomogeneous plate, where the effective elastic modulus $E(\xi)$, effective Poisson's ratio $\nu(\xi)$, and effective mass density $\rho(\xi)$ are random fields. They depend on two principal sources of uncertainties: (1) randomness in the volume fraction of SiC particles $\phi_{\text{SiC}}(\xi)$, which varies only along ξ , and (2) randomness in constituent material properties, comprising elastic moduli E_{SiC} and E_{Al} , Poisson's ratios ν_{SiC} and ν_{Al} , and mass densities ρ_{SiC} and ρ_{Al} of SiC and Al material phases, respectively. The particle volume fraction $\phi_{\text{SiC}}(\xi)$ is a one-dimensional, inhomogeneous, Beta random field with mean $\mu_{\text{SiC}}(\xi) = 1 - \xi/L$, standard deviation $\sigma_{\text{SiC}}(\xi) = (\xi/L)(1 - \xi/L)$, where L is the length of the plate. Assuming an appropriately bounded covariance function of $\phi_{\text{SiC}}(\xi)$, the standardized volume fraction, $\tilde{\phi}_{\text{SiC}}(\xi) := [\phi_{\text{SiC}}(\xi) - \mu_{\text{SiC}}(\xi)]/\sigma_{\text{SiC}}(\xi)$, was mapped to a zero-mean, homogeneous, Gaussian image field $\alpha(\xi)$ with an exponential covariance function $\Gamma_{\alpha}(t) := \mathbb{E}[\alpha(\xi)\alpha(\xi + t)] = \exp(-|t|/0.125L)$ via $\tilde{\phi}_{\text{SiC}}(\xi) = F_{\text{SiC}}^{-1}[\Phi(\alpha(\xi))]$, where Φ is the distribution function of a standard Gaussian random variable and F_{SiC} is the marginal distribution function of $\tilde{\phi}_{\text{SiC}}(\xi)$. The Karhunen-Loève approximation [5] was employed to discretize $\alpha(\xi)$ and hence $\phi_{\text{SiC}}(\xi)$ into 28 standard Gaussian random variables. In addition, the constituent material properties, E_{SiC} , E_{Al} , ν_{SiC} , ν_{Al} , ρ_{SiC} , and ρ_{Al} , were modeled as independent lognormal random variables with their means and coefficients of variation described in Table 3. Therefore, a total of 34 random variables are involved in this example. Employing a rule of mixture, $E(\xi) \cong E_{\text{SiC}}\phi_{\text{SiC}}(\xi) + E_{\text{Al}}[1 - \phi_{\text{SiC}}(\xi)]$, $\nu(\xi) \cong \nu_{\text{SiC}}\phi_{\text{SiC}}(\xi) + \nu_{\text{Al}}[1 - \phi_{\text{SiC}}(\xi)]$, and $\rho(\xi) \cong \rho_{\text{SiC}}\phi_{\text{SiC}}(\xi) + \rho_{\text{Al}}[1 - \phi_{\text{SiC}}(\xi)]$. Using these spatially-variant effective properties, a 20×40 mesh consisting of 800 eight-noded, second-order shell elements, shown in Figure 7(b), was constructed for FEA, to determine the natural frequencies of the FGM plate. No damping was included. A Lanczos algorithm [3] was employed for calculating the eigenvalues.

The probability distributions of the first six natural frequencies of the functionally graded material plate were eval-

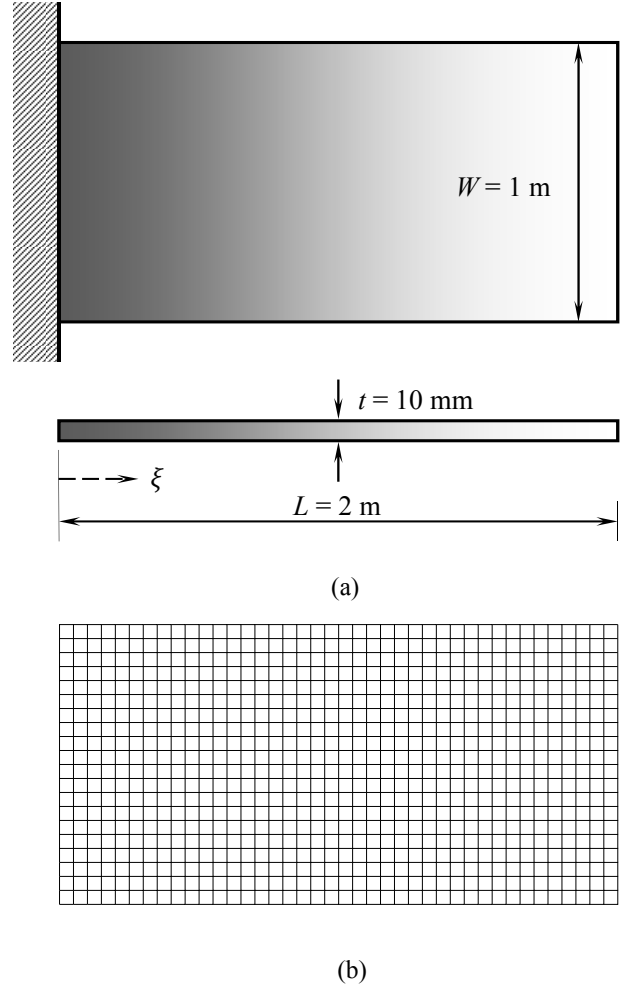


Figure 7: An FGM cantilever plate: (a) geometry; (b) a 20×40 FEA mesh.

Table 3: Statistical material properties of constituents in SiC-Al FGM.

Material properties ⁽¹⁾	Mean	COV ⁽²⁾ , %
E_{SiC} , GPa	419.2	15
ν_{SiC}	0.19	5
ρ_{SiC} , kg/m ³	3210	15
E_{Al} , GPa	69.7	15
ν_{Al}	0.34	5
ρ_{Al} , kg/m ³	2520	15

⁽¹⁾ E_{SiC} = elastic modulus of SiC, ν_{SiC} = Poisson's ratio of SiC, ρ_{SiC} = mass density of SiC, E_{Al} = elastic modulus of Al, ν_{Al} = Poisson's ratio of Al, ρ_{Al} = mass density of Al.

⁽²⁾ Coefficient of variation.

²Functionally graded materials are two- or multi-phase particulate composites in which material composition and microstructure vary spatially in the macroscopic length scale to meet a desired functional performance.

uated using four different PDD methods: (1) the bivariate partially adaptive-sparse PDD method with full-grid dimension-reduction integration; (2) the bivariate partially adaptive-sparse PDD method with sparse-grid dimension-reduction integration with extended FSI rule; (3) the univariate, fifth-order PDD method; and (4) the bivariate, fifth-order PDD method; and the crude MCS. Again, the order of the truncated PDD was selected based on the largest value of m_u required in the adaptive-sparse PDD methods. The tolerances used for adaptive and ranking algorithms are $\epsilon_1 = \epsilon_2 = 10^{-6}$ and $\epsilon_3 = 0.9$. Figure 8 presents the marginal probability distributions $F_i(\omega_i) := P[\Omega_i \leq \omega_i]$ of the first six natural frequencies $\Omega_i, i = 1, \dots, 6$, where all the PDD solutions were obtained from the embedded MCS. The plots are made over a semi-logarithmic scale to delineate the distributions in the tail regions. For all six frequencies, the probability distributions obtained from a bivariate partially adaptive-sparse PDD method, whether using either full-grid or sparse-grid, and the bivariate fifth-order PDD method are much closer to the crude Monte Carlo results compared with those obtained from the univariate, fifth-order PDD method. While all PDD approximations require fewer function evaluations than the crude MCS, both variants of the partially adaptive-sparse PDD approximations remit exceptionally high efficiency by an average factor of six when compared with the bivariate, fifth-order PDD approximation. However, the advantage of the sparse-grid integration over the full-grid integration employed in the adaptive-sparse approximation is modest in terms of computational efficiency. This is explained as follows.

The efficient reduced ranking algorithm was employed in this example. When the bivariate component functions were ranked for $m_u = 1$, the coefficient calculation for both full-grid and sparse-grid involved function evaluation at the point $(0, 0)$ as shown for $l = 1$ in Figure 2. The function evaluations at this point return only the functions already evaluated at the point (c) , i.e., response at mean $y(c)$, thus the bivariate component functions could not be ranked for $m_u = 1$. When the polynomial order was incremented to $m_u = 2$, the full-grid for $l = 2$ comprises of four non-zero integration points, resulting in non-trivial bivariate function evaluations at those points. However, the sparse-grid consists of four new points lying on the axes, failing to capture the interaction effect of two variables. This results in bivariate function evaluations that are not useful in creating a ranking. Thus, for $m_u = 2$, full-grid involves ranking all the $28 \times 27/2 = 378$ bivariate component functions, with $378 \times 4 = 1512$ new function evaluations, while the sparse-grid was still lacking any ranking. Moving to $m_u = 3$, full-grid can afford to exploit the efficient reduced-ranking by truncating the ranking from $m_u = 2$ and calculating coefficients only for fewer than 378 component functions. However, the sparse-grid is forced to evaluate all 378 component functions for $m_u = 3$, resulting in $378 \times 4 = 1512$ function evaluations at four new integration points, depriving this efficient technique of any initial advantage. The modest advantage in computational efficiency that the sparse-grid eventually achieves was obtained only after ranking at $m_u = 4$ and onwards.

Figure 9 displays the joint probability density function $f_{12}(\omega_1, \omega_2)$ of the first two natural frequencies Ω_1 and Ω_2 ob-

tained by the two variants of the bivariate partially adaptive-sparse PDD method, the bivariate, fifth-order PDD method, and crude MCS. Although visually comparing these three-dimensional plots is not simple, the joint distributions from all PDD approximations and the crude Monte Carlo method seem to match reasonably well. The contours of these three-dimensional plots were studied at two notably different levels: $f_{12} = 0.005$ (high level) and $f_{12} = 0.0005$ (low level), as depicted in Figures 10(a) and 10(b), respectively. For both levels examined, a good agreement exists among the contours from all four distributions. These results are consistent with the marginal distributions of natural frequencies discussed in the preceding paragraph.

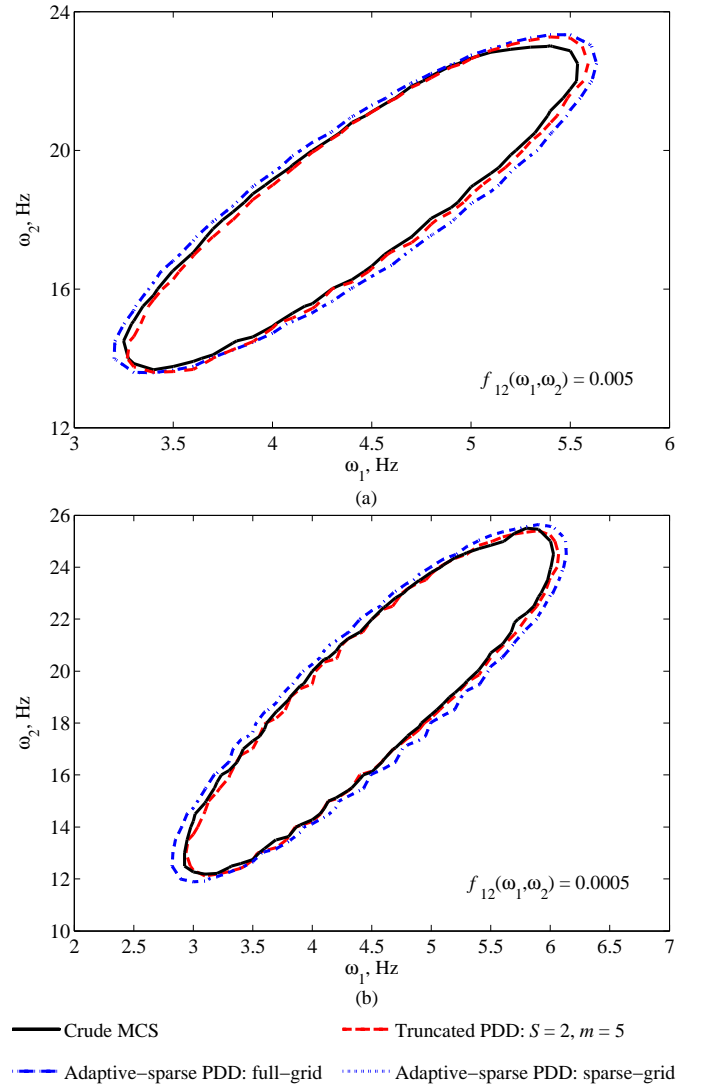


Figure 10: Contours of the joint density function of the first and second natural frequencies of the FGM plate by various PDD approximations and crude MCS: (a) $f_{12} = 0.005$; (b) $f_{12} = 0.0005$.

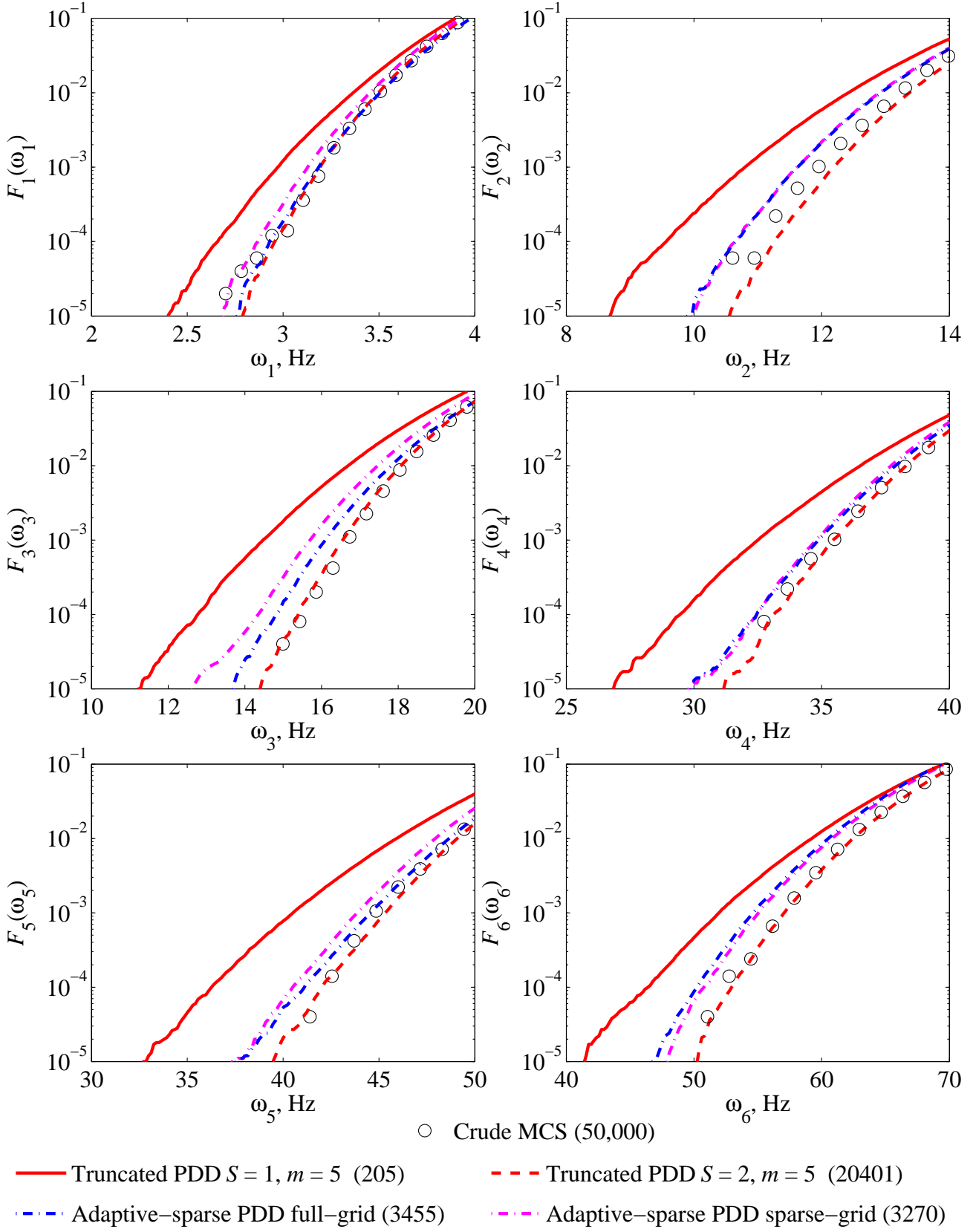


Figure 8: Marginal probability distributions of the first six natural frequencies of an FGM plate by various PDD approximations and crude MCS.

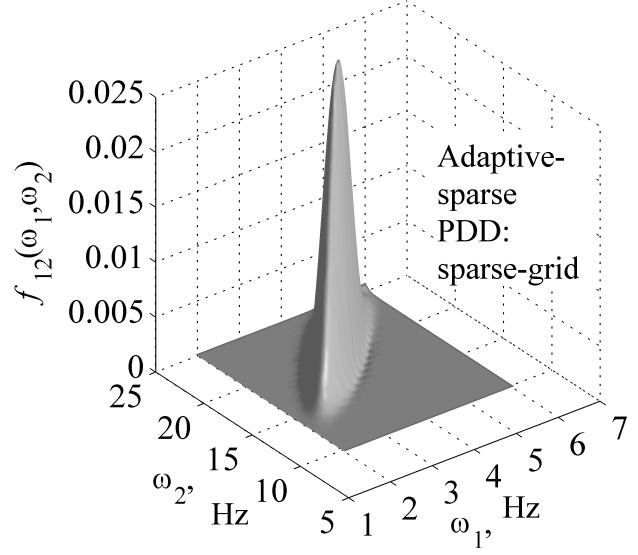
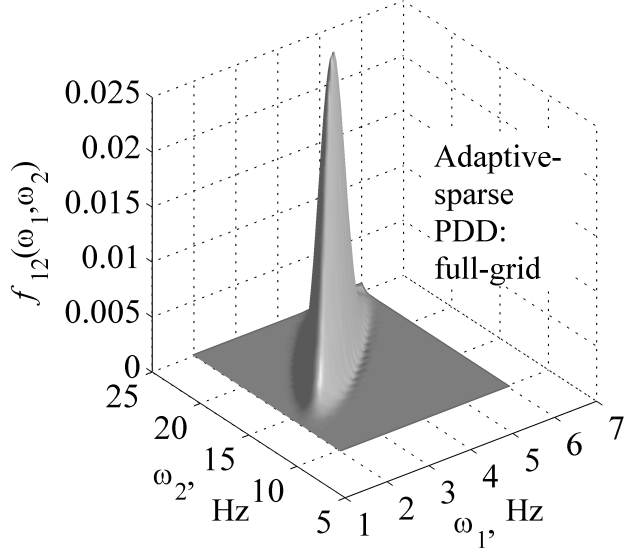
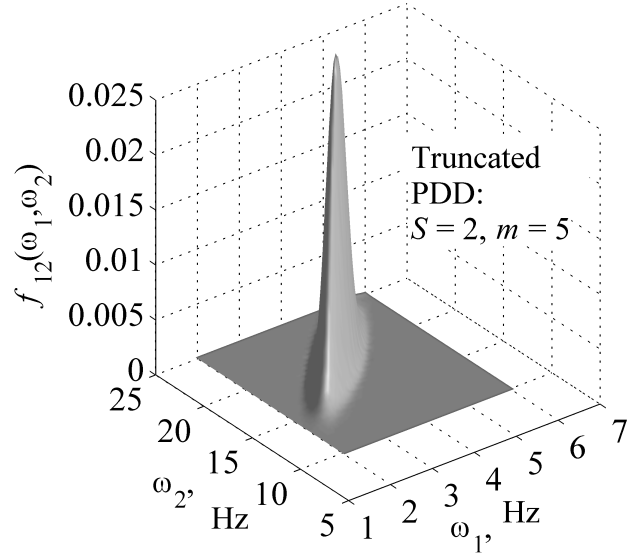
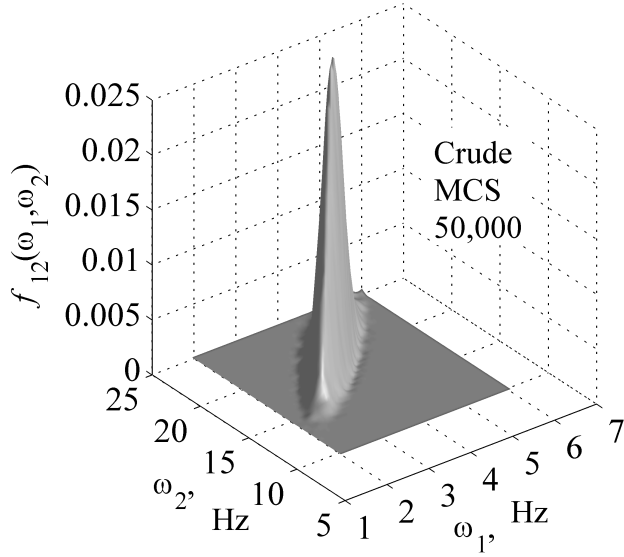


Figure 9: Joint probability density function of the first and second natural frequencies of the FGM plate by various PDD approximations and crude MCS.

6. Application: A Disk Brake System

This section demonstrates the capabilities of the proposed partially adaptive-sparse PDD method in solving a large-scale practical engineering problem. The application comprises of determining instabilities in a disk brake system in terms of statistical analysis of complex frequencies and corresponding mode shapes. The dynamic instabilities in a braking system, emanating from complex frequencies, give rise to the highly undesired phenomenon of brake squeal. When a braking system is subjected to random input parameters, it is imperative to perform a random brake-squeal analysis in order to identify, quantify, and minimize the random dynamic instabilities.

6.1. Brake-Squeal Analysis

A disk brake system, illustrated in Figure 11(a), slows motion of the wheel by pushing brake pads against a rotor with a set of calipers [7]. The brake pads mounted on a brake caliper is forced mechanically, hydraulically, pneumatically, or electromagnetically against both sides of the rotor. Friction causes the rotor and attached wheel to slow or stop. Figure 11(b) presents a simplified FEA model of a disk brake system commonly used in domestic passenger vehicles. The system consists of a rotor of diameter 288 mm and thickness 20 mm. Two pads are positioned on both sides of the rotor. Assembled behind the pads are back plates and insulators. The FEA mesh of the model consists of 26,125 elements and 111,129 active degrees of freedom and was generated using C3D6 and C3D8I elements in Abaqus computer software (Version 6.12) [4]. The rotor is made of cast iron and the back plates and insulators are made of steel. The two brake pads are made of organic frictional material, which is modeled as an orthotropic elastic material. The mass densities and Young's moduli of the rotor, back-plates, insulators and pads along with the shear moduli of the pads are modeled as random variables with uniform distribution. Along with the random material properties, the brake pressure, the radial velocity of the rotor, and the coefficient of friction between the rotor and pads are modeled as uniform random variables, constituting a total of 16 random variables in this problem. The statistical properties of all random variables are listed in Table 4. Apart from the random material properties, the deterministic Poisson's ratio of rotor, back-plates and insulators are 0.24, 0.28, and 0.29, respectively. The three Poisson's ratios of orthotropic material of pads are $\nu_{12} = 0.06$, $\nu_{23} = 0.41$, and $\nu_{31} = 0.15$.

6.2. Results

The dynamic analysis was performed in four steps. In the first step, contact was established between the rotor and the pad by applying brake pressure to the external surfaces of the insulators. Braking at low velocity was simulated in the second step by imposing a rotational velocity on the rotor, accompanied with an introduction of a non-zero friction coefficient between rotor and pad. In the third step, natural frequencies up to 20 kHz were extracted by the eigenvalue extraction procedure in the steady-state condition using the automatic multilevel substructuring method with subspace projection in Abaqus. Finally, in

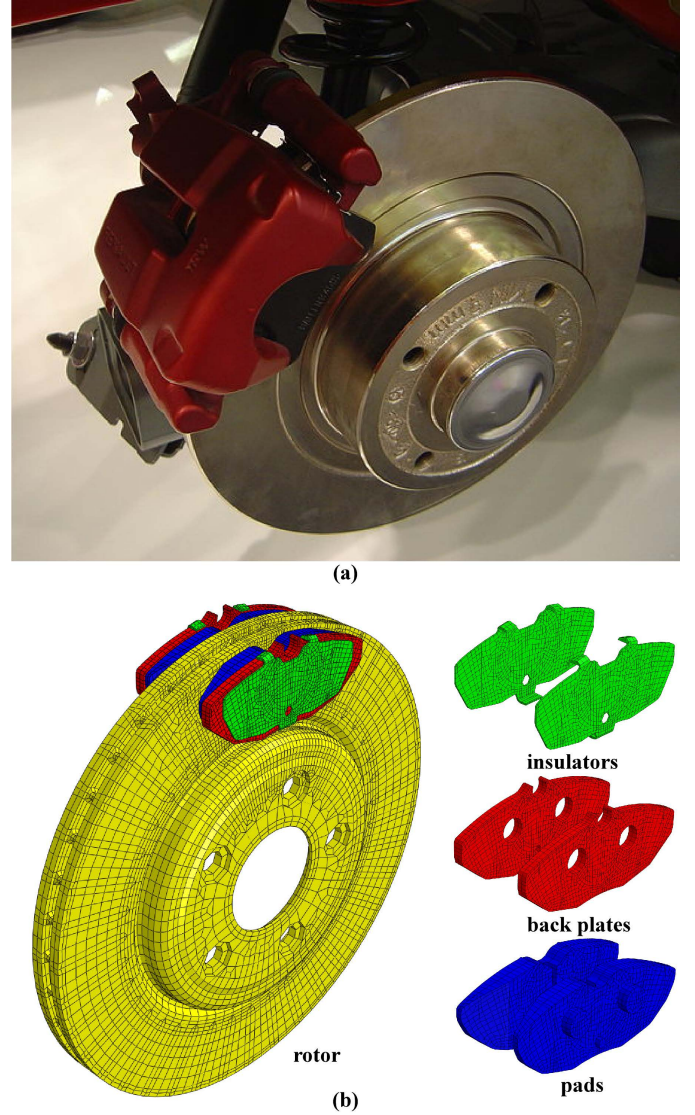


Figure 11: A disk brake system with various mechanical components: (a) close-up on a passenger vehicle; (b) a simplified FEA model

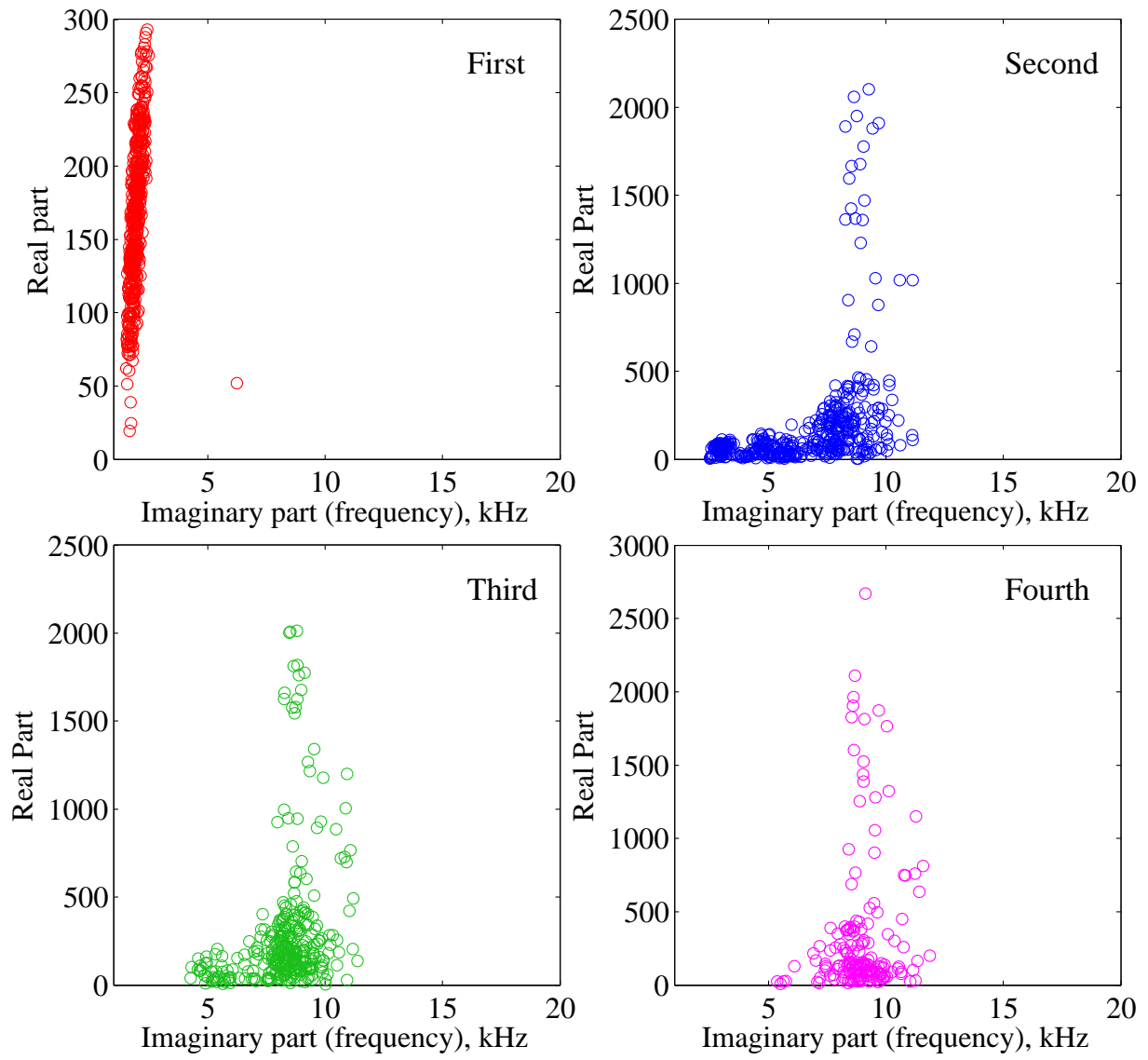


Figure 12: Complex eigenvalues of a disk brake system for first four unstable modes.

Table 4: Random input variables in disk-brake system with the minimum (a_i) and maximum (b_i) values of their uniform distributions.

Random variables ⁽¹⁾	a_i	b_i
ρ_{rotor} , kg/mm ³	5.329×10^{-6}	9.071×10^{-6}
$\rho_{\text{back plate}}$, kg/mm ³	5.788×10^{-6}	9.851×10^{-6}
$\rho_{\text{insulator}}$, kg/mm ³	5.788×10^{-6}	9.851×10^{-6}
ρ_{pad} , kg/mm ³	1.858×10^{-6}	3.162×10^{-6}
E_{rotor} , GPa	92.52	157.5
$E_{\text{back plate}}$, GPa	153.2	260.8
$E_{\text{insulator}}$, GPa	153.2	260.8
$E_{1,\text{pad}}$, GPa	4.068	6.924
$E_{2,\text{pad}}$, GPa	4.068	6.924
$E_{3,\text{pad}}$, GPa	1.468	2.498
$G_{12,\text{pad}}$, GPa	1.917	3.263
$G_{13,\text{pad}}$, GPa	0.873	1.486
$G_{23,\text{pad}}$, GPa	0.873	1.486
P , kg/mm ²	370.1	629.9
ω , rad/s	3.701	6.299
μ	0.50	0.70

(1) ρ_{rotor} , $\rho_{\text{back plate}}$, $\rho_{\text{insulator}}$, ρ_{pad} : mass densities of corresponding materials, E_{rotor} , $E_{\text{back plate}}$, $E_{\text{insulator}}$: elastic modulus of corresponding materials, $E_{1,\text{pad}}$, $E_{2,\text{pad}}$, $E_{3,\text{pad}}$: elastic modulus associated with the normal directions of pad material, $G_{12,\text{pad}}$, $G_{13,\text{pad}}$, $G_{23,\text{pad}}$: shear modulus associated with the principal directions of pad material, P : brake pressure, ω : radial velocity, μ : friction coefficient.

the fourth step a complex eigenvalue analysis was performed up to the first 55 modes.

The bivariate partially adaptive-sparse PDD method with tolerances $\epsilon_1 = \epsilon_2 = 10^{-6}$, $\epsilon_3 = 0.9$ was applied to determine the probabilistic characteristics of the dynamic instabilities caused by the first two unstable modes of the disk brake system. Since all input random variables are uniformly distributed, classical Legendre orthonormal polynomials were used as basis functions. The PDD coefficients were calculated using the quasi MCS with 500 samples generated from a 16-dimensional low-discrepancy Sobol sequence. The sample size, although selected arbitrarily, is adequate, as there exist no significant changes to the coefficients, at least, for this problem. Figure 12 displays real and imaginary parts of the eigenvalues of the first four unstable modes obtained in each quasi Monte Carlo sample. These unstable modes, conveyed by complex frequencies with positive real parts, reflect the dynamic instability caused in the brake system. Each occurrence of the unstable frequency may cause the brake to squeal.

Equations (29) and (31) were employed to calculate the second-moment statistics of each nodal displacement component of an eigenvector describing the associated mode shape of the disk brake system. Based on these statistics, the \mathcal{L}_2 -norms, that is, the square root of sum of squares, of the mean and variance of a nodal displacement were calculated. Figures 13(a) and 13(b) present contour plots of the \mathcal{L}_2 -norms of the means and variances, respectively, of the first two unstable mode shapes of the disk brake system. Similar results can be generated for other mode shapes, stable or unstable, if desired.

For a disk brake system with complex frequencies, the i th ef-

fective damping ratio is defined as $-2\text{Re}[\lambda_u^{(i)}(\mathbf{X})]/\text{Im}[\lambda_u^{(i)}(\mathbf{X})]$, where $\text{Re}[\lambda_u^{(i)}(\mathbf{X})]$ and $\text{Im}[\lambda_u^{(i)}(\mathbf{X})]$ are the real part and the imaginary part, respectively, of the i th unstable frequency $\lambda_u^{(i)}(\mathbf{X})$. The magnitude of the damping ratio represents the harshness of brake squeal. The larger the magnitude of the damping ratio, the higher the propensity for brake squeal. Figure 14 illustrates the marginal probability density functions of the effective damping ratios corresponding to the first two unstable modes. These probability densities provide a measure of the effect of random input parameters on the dynamic instabilities caused in the disk brake system.

It is worth mentioning that a similar brake-squeal analysis with only five input random variables was performed using a univariate RDD method [24]. However, verification or improvement of the univariate solution was not possible due to inherent limitations of the method used. The adaptive-sparse PDD approximations developed in this work have overcome this quandary even for a significantly more input variables.

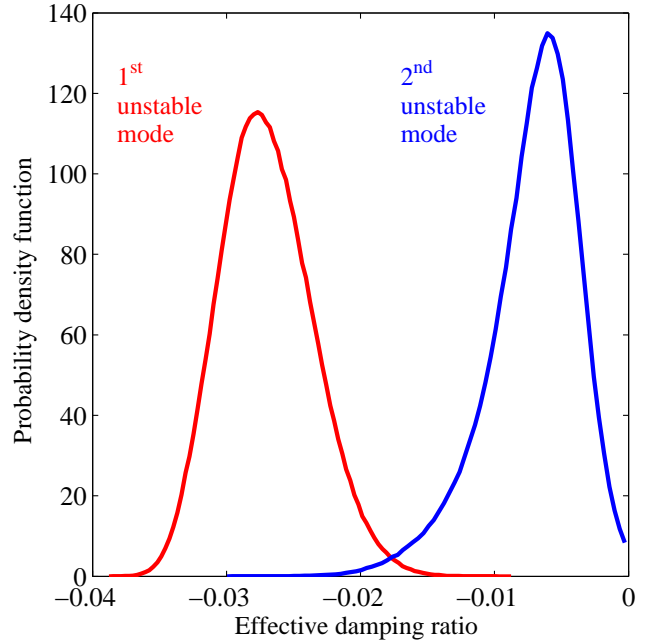
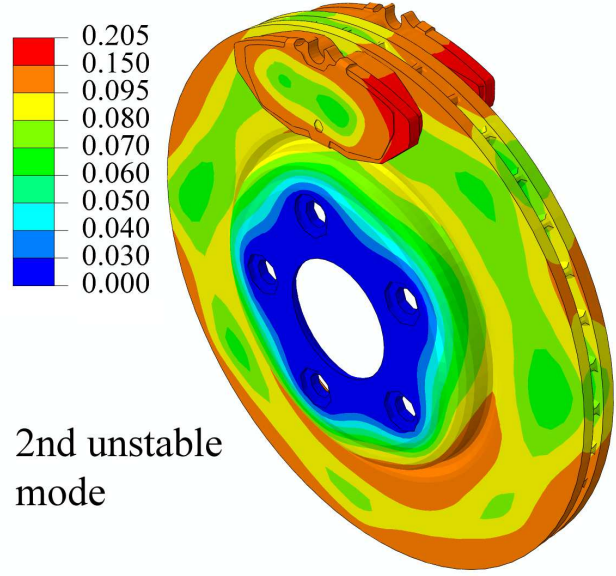
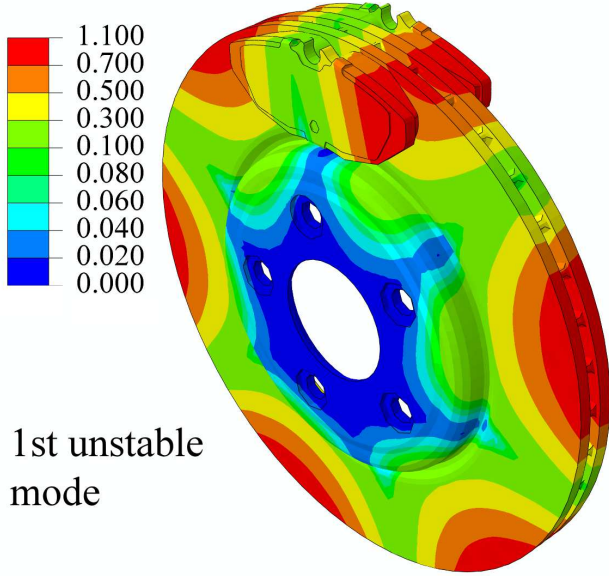


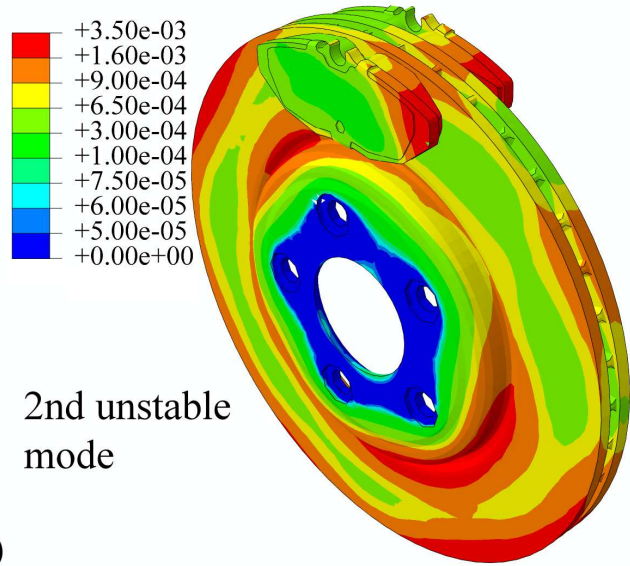
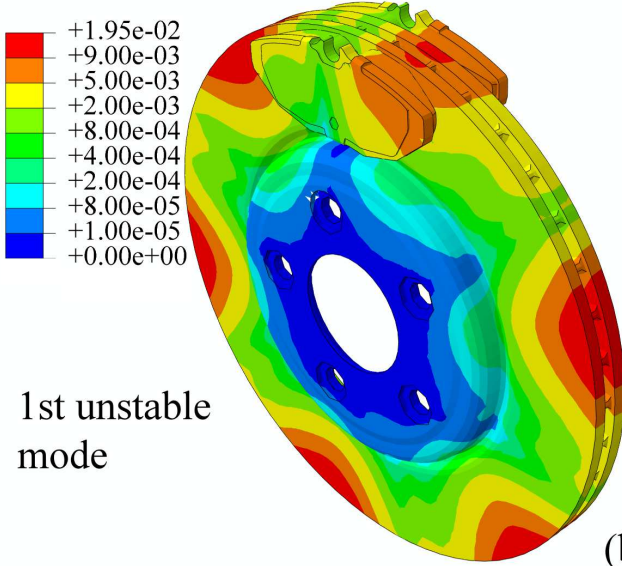
Figure 14: Marginal probability density functions of the effective damping ratios of first two unstable modes of a disk brake system by the bivariate partially adaptive-sparse PDD method.

7. Conclusions

Two new adaptive-sparse PDD methods, the fully adaptive-sparse PDD method and a partially adaptive-sparse PDD method, were developed for uncertainty quantification of high-dimensional complex systems commonly encountered in applied sciences and engineering. The methods are based on global sensitivity analysis for defining the pruning criteria to retain important PDD component functions, and a full- or sparse-grid dimension-reduction integration or quasi MCS for estimating the PDD expansion coefficients. In the fully adaptive-sparse



(a)



(b)

Figure 13: Contour plots of the \mathcal{L}_2 -norm of the first two unstable mode shapes of a disk brake system by the bivariate partially adaptive-sparse PDD method: (a) mean; (b) variance.

PDD approximation, PDD component functions of an arbitrary number of input variables are retained by truncating the degree of interaction among input variables and the order of orthogonal polynomials according to specified tolerance criteria. In a partially adaptive-sparse PDD approximation, PDD component functions with a specified degree of interaction are retained by truncating the order of orthogonal polynomials, fulfilling relaxed tolerance criteria. The former approximation is comprehensive and rigorous, leading to the second-moment statistics of a stochastic response that converges to the exact solution when the tolerances vanish. The latter approximation, obtained through regulated adaptivity and sparsity, is more economical than the former approximation and is, therefore, expected to solve practical problems with numerous variables. A unified computational algorithm was created for solving a general stochastic problem by the new PDD methods. Two distinct ranking schemes – full ranking and reduced ranking – were also developed for grading PDD component functions in the unified algorithm. Compared with past developments, the adaptive-sparse PDD methods do not require truncation parameter(s) to be assigned a priori or arbitrarily. In addition, two numerical techniques, one employing a nested sparse-grid dimension-reduction integration and the other exploiting quasi MCS, were applied for the first time to estimate the PDD expansion coefficients both accurately and efficiently.

The adaptive-sparse PDD methods were employed to calculate the second-moment properties and tail probability distributions in three numerical problems, where the output functions are either simple mathematical functions or eigenvalues of dynamic systems, including natural frequencies of a three-degree-of-freedom linear oscillator and an FGM plate. The mathematical example reveals that the user-defined tolerances of an adaptive-sparse PDD method are closely related to the relative error in calculating the variance, thus providing an effective tool for modulating the accuracy of the resultant approximation desired. Since the adaptive-sparse PDD approximation retains only the component functions with significant contributions, it is also able to achieve a desired level of accuracy with considerably fewer coefficients than required by existing truncated PDD approximations. The results of the linear oscillator display a distinct advantage of the reduced ranking system over the full ranking system, as the former requires significantly fewer expansion coefficients to achieve results nearly identical to those of the latter. For a required level of accuracy in calculating the tail probabilistic characteristics of natural frequencies of an FGM plate, the new bivariate adaptive-sparse PDD method is more economical than the existing bivariate truncated PDD method by almost an order of magnitude. Finally, the new PDD method was successfully applied to solve a stochastic dynamic instability problem in a disk brake system, demonstrating the ability of the new methods in handling industrial-scale problems.

References

- [1] R. Bellman. *Dynamic Programming*. Princeton University Press: Princeton, NJ, 1957.
- [2] G. Blatman and B. Sudret. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. *Comptes Rendus Mcanique*, 336(6):518 – 523, 2008.
- [3] J. K. Cullum and R. A. Willoughby. *Lanczos Algorithms for Large Symmetric Eigenvalue Computations: Theory*. Classics in applied mathematics. Society for Industrial and Applied Mathematics, 2002.
- [4] Dassault Systems Simulia Corp. *Abaqus Standard, Version 6.11*, 2011.
- [5] W. B. Davenport and W. L. Root. *An Introduction to the Theory of Random Signals and Noise*. McGraw-Hill, New York, NY, 1958.
- [6] M. K. Deb, I. M. Babuška, and J. T. Oden. Solution of stochastic partial differential equations using galerkin finite element techniques. *Computer Methods in Applied Mechanics and Engineering*, 190(48):6359 – 6372, 2001.
- [7] Disk brake. In *Wikipedia*. Retrieved April, 2013.
- [8] B. Efron and C. Stein. The jackknife estimate of variance. *The Annals of Statistics*, 9(3):pp. 586–596, 1981.
- [9] H. Faure. Discrepances de suites associées à un système de numération (en dimension un). *Bulletin de la Société Mathématique de France*, 109:143–182, 1981.
- [10] B. Ganapathysubramanian and N. Zabaras. Sparse grid collocation schemes for stochastic natural convection problems. *Journal of Computational Physics*, 225(1):652 – 685, 2007.
- [11] W. Gautschi. *Orthogonal polynomials: computation and approximation*. Numerical mathematics and scientific computation. Oxford University Press, 2004.
- [12] A. Genz. Fully symmetric interpolatory rules for multiple integrals. *SIAM Journal on Numerical Analysis*, 23(6):pp. 1273–1283, 1986.
- [13] A. Genz and B. D. Keister. Fully symmetric interpolatory rules for multiple integrals over infinite regions with gaussian weight. *Journal of Computational and Applied Mathematics*, 71(2):299 – 309, 1996.
- [14] R. Ghanem and P. D. Spanos. *Stochastic finite elements: a spectral approach*. World Publishing Corp., 1991.
- [15] M. Griebel and M. Holtz. Dimension-wise integration of high-dimensional functions with applications to finance. *J. Complex.*, 26(5):455–489, October 2010.
- [16] J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2(1):84–90, 1960.
- [17] F. J. Hickernell. Quadrature error bounds with applications to lattice rules. *SIAM J. Numer. Anal.*, 33:1995–2016, 1996.
- [18] M. Holtz. *Sparse grid quadrature in high dimensions with applications in finance and insurance*. Ph.D. Dissertation, Bonn, Germany, 2008.
- [19] R. Li and R. Ghanem. Adaptive polynomial chaos expansions applied to statistics of extremes in nonlinear random vibration. *Probabilistic Engineering Mechanics*, 13(2):125 – 136, 1998.
- [20] X. Ma and N. Zabaras. An adaptive high-dimensional stochastic model representation technique for the solution of stochastic partial differential equations. *Journal of Computational Physics*, 229(10):3884 – 3915, 2010.
- [21] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1992.
- [22] E. Novak and K. Ritter. Simple cubature formulas with high polynomial exactness. *Constructive Approximation*, 15(4):499–522, 1999.
- [23] H. Rabitz and O. Alis. General foundations of high dimensional model representations. *Journal of Mathematical Chemistry*, 25:197–233, 1999. 10.1023/A:1019188517934.
- [24] S. Rahman. Stochastic dynamic systems with complex-valued eigen-solutions. *International Journal for Numerical Methods in Engineering*, 71:963–986, 2007.
- [25] S. Rahman. A polynomial dimensional decomposition for stochastic computing. *International Journal for Numerical Methods in Engineering*, 76:2091–2116, 2008.
- [26] S. Rahman. Extended polynomial dimensional decomposition for arbitrary probability distributions. *Journal of Engineering Mechanics*, 135(12):1439–1451, 2009.
- [27] S. Rahman. Statistical moments of polynomial dimensional decomposition. *Journal of Engineering Mechanics*, 136(7):923–927, 2010.
- [28] S. Rahman. Decomposition methods for structural reliability analysis revisited. *Probabilistic Engineering Mechanics*, 26(2):357 – 363, 2011.
- [29] S. Rahman. Global sensitivity analysis by polynomial dimensional de-

- composition. *Reliability Engineering & System Safety*, 96(7):825 – 837, 2011.
- [30] S. Rahman. Approximation errors in truncated dimensional decompositions. *Accepted in Mathematics of Computation*, 2013.
 - [31] S. Rahman and X. Ren. Novel computational method for high-dimensional stochastic sensitivity analysis. *Accepted in International Journal of Numerical Methods in Engineering*, 2013.
 - [32] S. Rahman and V. Yadav. Orthogonal polynomial expansions for solving random eigenvalue problems. *International Journal for Uncertainty Quantification*, 1:163–187, 2011.
 - [33] I.H. Sloan and S. Joe. *Lattice Methods for Multiple Integration*. Oxford science publications. Clarendon Press, 1994.
 - [34] I. M. Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *U.S.S.R. Comput. Math. Math. Phys.*, 7:86–112, 1967.
 - [35] I. M. Sobol. Theorems and examples on high dimensional model representation. *Reliability Engineering & System Safety*, 79(2):187 – 193, 2003.
 - [36] I.M. Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(13):271 – 280, 2001.
 - [37] X. Wan and G. E. Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *Journal of Computational Physics*, 209(2):617 – 642, 2005.
 - [38] X. Wang. Improving the rejection sampling method in quasi-monte carlo methods. *Journal of Computational and Applied Mathematics*, 114(2):231 – 246, 2000.
 - [39] D. Xiu and G. E. Karniadakis. The wiener-askey polynomial chaos for stochastic differential equations. *SIAM Journal of Scientific Computing*, 24:619–644, 2002.
 - [40] H. Xu and S. Rahman. A generalized dimension-reduction method for multi-dimensional integration in stochastic mechanics. *International Journal for Numerical Methods in Engineering*, 61:1992–2019, 2004.
 - [41] H. Xu and S. Rahman. Decomposition methods for structural reliability analysis. *Probabilistic Engineering Mechanics*, 20(3):239 – 250, 2005.
 - [42] V. Yadav and S. Rahman. Uncertainty quantification of high-dimensional complex systems by multiplicative polynomial dimensional decompositions. *International Journal for Numerical Methods in Engineering*, 94(3):221–247, 2013.
 - [43] X. Yang, M. Choi, G. Lin, and G.E. Karniadakis. Adaptive anova decomposition of stochastic incompressible and compressible flows. *Journal of Computational Physics*, 231(4):1587 – 1614, 2012.